

Handleiding Thema04

Ronald Wedema & Arne Poortinga

Inleiding

Computer programma's worden gebruikt om gegevens te verwerken. Binnen de bio-informatica gaat het dan niet over willekeurige gegevens, maar specifiek om gegevens die bedoeld zijn om een bepaalde onderzoeksvraag te beantwoorden. Klinkt goed, maar wat is een (goede) onderzoeksvraag eigenlijk? Kun je met de beschikbare gegevens de onderzoeksvraag überhaupt wel beantwoorden, waar komen die gegevens vandaan en zijn de gegevens wel betrouwbaar? In dit thema en de ondersteunende vakken krijg je antwoord op deze vragen. Om ze gestructureerd aan je aan te bieden en niet als een ongestructureerde brei over de schutting te gooien, is de leerstof door ons opgedeeld in een aantal leerdoelen.

Leerdoelen

De leerdoelen voor deze themaopdracht staan als volgt in Osiris:

- De student is in staat om met een groep medestudenten data te verzamelen volgens een eenvoudig protocol.
- De student is in staat om zelfstandig eenvoudige data te verwerken, te analyseren en te visualiseren met behulp van R en RStudio.
- De student is in staat om met een groep medestudenten een eenvoudige materiaal en methoden sectie te schrijven.
- De (wikimedia.org) student is zelfstandig in staat om duidelijke figuren met correcte, duidelijke bijschriften te maken die het beantwoorden van een onderzoeksvraag ondersteunen.
- De student is in staat om zelfstandig een eenvoudige dataset statistisch te analyseren en de resultaten hiervan samen te vatten zodat een eenvoudige onderzoeksvraag beantwoord kan worden.
- De student laat uit zichzelf een professionele werkhouding zien.
- De student is zelfstandig in staat om eenvoudige feedback op een eenvoudig werkstuk te geven en te ontvangen.

De laatste twee leerdoelen hebben niet direct te maken met het doen van onderzoek of het verwerken van gegevens maar wel met het ontwikkelen van gedrag wat verwacht wordt van een professional.

Beoordeling

In onderstaande tabel staan voor elk leerdoel de criteria aangegeven waarop je beoordeeld wordt en welke puntenrange er bij bepaalde 'gedragskenmerken' hoort. Nadat je de opdrachten die horen bij een bepaald leerdoel volbracht hebt, zal je docent ze beoordelen en je score voor dit onderdeel en zijn motivatie hiervoor invullen. Je eindcijfer wordt berekend door de punten die je gehaald hebt bij elkaar op te tellen en te delen door het maximaal aantal te behalen punten (tien per leerdoel) en de uitkomst hiervan met tien te vermenigvuldigen. Je hebt een voldoende als je een 5.5 of hoger haalt. Omdat de feedback belangrijk is om je leerdoelen te halen, wordt je aangeraden om altijd mondelinge feedback van je docent te vragen. Mocht je voor een onderdeel laag gescoord hebben, neem dan *direct* contact met je docent op om de mogelijkheid tot het maken van een herkansing te bespreken.

Als je een opdracht niet inlevert of op een andere manier c.q. een ander formaat dan aangegeven wordt, dan krijg je 0 punten.

leerdoel	slecht/ onvoldoende	zwak/ voldoende	gemiddeld/ goed	excellent
	4 tot 6	6 tot 8	8 tot 10	10
De student is in staat om met een groep medestudenten een goede onderzoeksvraag te formuleren voor een eenvoudig onderzoek.	De opdracht is maar deels ingeleverd. De vraagstelling is overduidelijk niet ter zake doend, zeer slecht geformuleerd of moeilijk leesbaar door grammaticale fouten. Het ingeleverde werk is niet op spelling gecontroleerd.	De vraagstelling is niet duidelijk of niet eenduidig beantwoordbaar	De vraagstelling is eenvoudig, helder en eenduidig beantwoordbaar	De vraagstelling is duidelijk, eenduidig beantwoordbaar en prikkelend
De student is	De opdracht is maar deels	Het protocol is niet	Het protocol maakt het	Het protocol

in staat om met een groep medestudenten een eenvoudige, duidelijk protocol te schrijven om data te beantwoording van een eenvoudige onderzoeksvraag te verzamelen.	ingeleverd. Het protocol is overduidelijk niet ter zake doend, zeer slecht geformuleerd of moeilijk leesbaar door grammaticale fouten. Het ingeleverde werk is niet op spelling gecontroleerd.	duidelijk of niet toereikend c.q. ter zake doend voor het beantwoorden van de vraag	beantwoord en van de vraag mogelijk	maakt het beantwoorden van de vraag mogelijk en houdt rekening met alternatieven.
De student is in staat om met een groep medestudenten data te verzamelen volgens een eenvoudige protocol	De opdracht is maar deels ingeleverd. Het ingeleverde werk is zeer slecht geformuleerd of moeilijk leesbaar door grammaticale fouten. Het ingeleverde werk is niet op spelling gecontroleerd. Het protocol is niet gevolgd.	Het protocol is slecht gevolgd en/of afwijkingen zijn slecht of niet gedocumenteerd.	het protocol is grotendeels gevolgd, afwijkingen zijn duidelijk gedocumenteerd	het protocol is gevolgd en dit is aan te tonen m.b.v. een duidelijke log.
De student is	De opdracht is maar deels	Code ontbreekt	alle code is ingeleverd	Alle code is ingeleverd

in staat om zelfstandig eenvoudige data verwerken, te analyseren en te visualiseren met R en RStudio	ingeleverd. Het ingeleverde werk is zeer slecht geformuleerd of moeilijk leesbaar door grammaticale fouten. Het ingeleverde werk is niet op spelling gecontroleerd. Code ontbreekt grotendeels of werkt grotendeels niet. De student kan de ingeleverde code niet uitleggen.	deels c.q. de ingeleverde code werkt deels niet en/of de student kan de ingeleverde code niet goed uitleggen	en werkt en de student kan de werking van de code uitleggen	en werkt, de student kan de werking van de code uitleggen en de code is zodanig gedocumenteerd en ingedeeld dat ze makkelijk opnieuw te gebruiken is.
De student is in staat om samen met medestudenten een eenvoudige materiaal en methoden sectie te schrijven.	De opdracht is maar deels ingeleverd. Het ingeleverde werk is zeer slecht geformuleerd of moeilijk leesbaar door grammaticale fouten. Het ingeleverde werk is niet op spelling gecontroleerd. Het ingeleverde werk is niet of nauwelijks	De beschrijving is oppervlakkig en/of bemoeilijkt opnieuw uitvoeren van het onderzoek	De beschrijving is voldoende om het onderzoek te kunnen herhalen	De beschrijving is voldoende om het onderzoek te kunnen herhalen en de sectie voldoet ook qua vorm aan de gestelde eisen.

	van toepassing op het onderwerp.			
De student is zelfstandig in staat om duidelijke figuren met correcte, duidelijke bijschrift en te maken die het beantwoorden van een onderzoeksvraag ondersteunen	De opdracht is maar deels ingeleverd. Het ingeleverde werk is zeer slecht geformuleerd of moeilijk leesbaar door grammaticale fouten. Het ingeleverde werk is niet op spelling gecontroleerd.	De figuren zijn niet compleet, passen niet bij de data of geven een misleidend beeld, of de bijschriften verduidelijken en de grafiek niet, of de figuur draagt niet bij aan het beantwoorden van de onderzoeksvraag.	De figuren zijn compleet inclusief assen en legenda en de bijschriften beschrijven correct de grafiek en de figuren ondersteunen het beantwoorden van de onderzoeksvraag.	De figuren zijn compleet inclusief assen en legenda en zijn aantrekkelijk vormgegeven en geven een duidelijk beeld van de data. De bijschriften beschrijven correct de grafiek en de figuren ondersteunen het beantwoorden van de onderzoeksvraag.
De student is in staat om zelfstandig een eenvoudige dataset statistisch te analyseren en de resultaten hiervan samen te	De opdracht is maar deels ingeleverd. Het ingeleverde werk is zeer slecht geformuleerd of moeilijk leesbaar door grammaticale fouten. Het ingeleverde werk is niet op spelling gecontroleerd.	De analyse of de samenvatting zijn onvolledig of onjuist of dragen niet bij aan het beantwoorden van de onderzoeksvraag	De analyse en de samenvatting zijn volledig en juist en dragen bij aan het beantwoorden van de onderzoeksvraag	de analyse en de samenvatting zijn volledig en juist en zijn duidelijk, overzichtelijk en samenhangend beschreven en dragen bij aan het beantwoorden van de

vatten zodat een eenvoudige onderzoeksvraag beantwoord kan worden.	d. De analyse of de samenvatting zijn zeer onvolledig of zeer onjuist.			onderzoeksvraag.
De student laat uit zichzelf een professionele werkhouding zien	De student is niet punctueel of laat zich, ook na relevante feedback, tijdens de practica regelmatig door andere zaken (telefoon, social media etc., niet-relevante websites) afleiden	De student is soms niet punctueel of laat zich tijdens de practica regelmatig door andere zaken (telefoon, social media etc., niet-relevante websites) afleiden	De student is punctueel en is tijdens de practica daar mee bezig	de student is punctueel en is tijdens de practica daar mee bezig en levert, bijvoorbeeld in discussies of door het stellen van vragen, een positieve bijdrage aan het practicum.
De student is zelfstandig in staat om eenvoudige feedback op een eenvoudig werkstuk te geven en te ontvangen	De opdracht is maar deels ingeleverd. Het ingeleverde werk is zeer slecht geformuleerd of moeilijk leesbaar door grammaticale fouten. Het ingeleverde werk is niet op spelling gecontroleerd	De student geeft onjuiste of irrelevante feedback of verwerkt relevante feedback niet juist	de student geeft duidelijke, relevante en constructieve feedback en verwerkt dito feedback op eigen werk	de student geeft duidelijke, relevante en constructieve feedback die inzicht toont in de materie en anderen deelgenoot maakt in dit inzicht en

	d. De student geeft duidelijk onjuiste of duidelijk irrelevante feedback of negeert relevante feedback			verwerkt relevante feedback op eigen werk.
--	--	--	--	--

Onder andere omdat een aantal leerdoelen over je gedrag gaan en je docent dit alleen in tijdens de lessen kan beoordelen is, net zoals bij andere practica, *aanwezigheid verplicht*. Het kan natuurlijk voor komen dat je ziek bent of dat er een andere reden is waarom je niet aanwezig kunt zijn. Je moet je dan wel vóór het practicum begint per mail met opgave van de reden voor je afwezigheid afmelden bij je docent *en* bij het secretariaat van ILST. Ben je vaker dan drie maal *met* opgave van reden afwezig, dan wordt je practicum niet beoordeeld. Hetzelfde geldt als je vaker dan één maal afwezig bent zonder afmelding.

Planning

Het kwartaal bestaat uit zeven lesweken plus een tentamenperiode. In principe worden er voor dit practicum elke lesweek twee dagdelen ingeroosterd, maar omdat er in dit kwartaal ook allerlei feestdagen vallen moet hier soms van afgeweken worden. Per lesweek (of per blok van twee dagdelen) wordt de stof van één of meer leerdoelen aangeboden samen met opdrachten. De opdrachten zijn deels bedoeld als hulpmiddel om de leerstof te verduidelijken maar worden deels ook gebruikt om te beoordelen of je het bijbehorende leerdoel behaald hebt.

Hier onder wordt per lesweek aangegeven wat er behandeld wordt en wat je globaal in moet leveren om voor de behandelde leerdoelen beoordeeld te worden. Het inleveren doe je via de desbetreffende link op BlackBoard. Hier staan ook de geldende deadlines.

Let Op! Omdat het voor kan komen dat je meer of minder dan twee practicumblokken in een week hebt, kan het gebeuren dat je meer dan één opdracht in een week in moet leveren! De weeknummers in dit document zijn alleen indicatief. **Check *altijd* BlackBoard voor de deadlines!** Let er ook op dat de deadline voor jouw deelklas af kan wijken van die van de andere deelklas!

week	leerdoel	hoe getoetst
1	De student is in staat om met een groep medestudenten een goede onderzoeksvraag te formuleren voor een eenvoudig onderzoek.	Je formuleert je eigen onderzoeksvraag.
2	De student is in staat om met een groep medestudenten een eenvoudig, duidelijk protocol te schrijven om data ter beantwoording van een eenvoudige onderzoeksvraag te verzamelen.	Je schrijft je eigen protocol om data te verzamelen om een bepaalde onderzoeksvraag te beantwoorden.
2	De student is zelfstandig in staat om eenvoudige feedback op een eenvoudig werkstuk te geven en te ontvangen.	Je geeft feedback op een protocol dat door andere groep geschreven is en krijgt zelf ook feedback van anderen. Verder beoordeelt de docent gedurende het hele

		blok hoe je met zijn/haar feedback omgaat en of je hier zelf om vraagt als dit nodig is.
3	De student is in staat om met een groep medestudenten data te verzamelen volgens een eenvoudig protocol.	Je wordt gevraagd gegevens te verzamelen volgens een protocol en dit loggen.
3 e.v.	De student is in staat om zelfstandig eenvoudige data verwerken, te analyseren en te visualiseren met R en RStudio.	Omgang met R/RStudio is een basisvaardigheid. Veel opdrachten die je maakt om aan te tonen dat je andere leerdoelen behaald hebt vereisen impliciet dat je goed hiermee om kunt gaan. Daarom wordt dit niet apart getoetst, maar wordt de kwaliteit van de code die je voor andere opdrachten inlevert gebruikt om dit leerdoel te toetsen. N.B. Het gebruik van RStudio en R-markdown om documenten te schrijven vindt plaats vanaf week 1.
4	De student is in staat om samen met medestudenten een eenvoudige materiaal en methoden sectie te schrijven.	Je schrijft een eigen Materiaal en Methoden sectie aan de hand van een gegeven en uitgevoerd protocol.
5	De student is zelfstandig in staat om duidelijke figuren met correcte, duidelijke bijschriften te maken die het beantwoorden van een onderzoeksvraag ondersteunen.	Je wordt gevraagd om aan de hand van een dataset en een onderzoeksvraag figuren te maken en in te leveren.
6/7	De student is in staat om zelfstandig een eenvoudige dataset statistisch te analyseren	Je maakt hiervoor een grotere opdracht.

	en de resultaten hiervan samen te vatten zodat een eenvoudige onderzoeksvraag beantwoord kan worden.	
voortdu rend	De student laat uit zichzelf een professionele werkhouding zien.	Tijdens de practicumuren wordt dit door je docent beoordeeld.

Zoals je in de [vakomschrijving](#) van dit thema kunt lezen krijg je, als je minimaal een 5,5 haalt tenminste, vier EC. Één EC staat voor ongeveer 27 uur verwacht werk. Dit betekent dus dat de totale studiebelasting voor dit vak 108 uur bedraagt in zeven weken. Per week dus ongeveer zestien uur. Per week zijn er ook nominaal acht contacturen ingeroosterd. Ga er dus van uit dat je buiten de contacturen om nog acht uur per week aan je thema werkt.

Week 1

Inleiding

Zoals je al in de tabel in het vorige hoofdstuk had gezien leer je in deze eerste week (de eerste twee practicumblokken) om een onderzoeksvraag te stellen en hoe je RStudio kunt gebruiken voor het schrijven van documenten. We beginnen met het laatste.

RStudio

In het vak [Informatica 4](#) maak je dit kwartaal onder andere kennis met [R](#), een programmeeromgeving die gericht is op het verkennen van datasets en het genereren van figuren op basis van die datasets. Hoewel je R-scripts net als Python scripts prima in een gewone teksteditor zoals [VIM](#) of [Geany](#) kunt schrijven heeft R z'n eigen IDE gekregen, [RStudio](#). Behalve dat RStudio het programmeren in R makkelijker maakt, biedt het je ook de mogelijkheid om *uitvoerbare code* en resultaten van die code te mixen met beschrijvende tekst. Dit is een enorm voordeel als je een dataset 'gaat verkennen'. Je kunt een log bijhouden van wat je precies gedaan hebt en je resultaten zonder dat je voortdurend tussen een tekstverwerkingsprogramma en een IDE c.q. een terminal heen en weer moet schakelen.

Markdown

Goede opmaak kan enorm helpen om een tekst makkelijk leesbaar te maken. Kale tekst blijkt hiervoor te beperkt. Daarom zijn er allerlei tekstverwerkers zoals LibreOffice Writer ontstaan die de gebruiker mogelijkheden bieden om tekst op te maken. Een alternatief voor het gebruik van gespecialiseerde tekstverwerkers is het gebruik van html. Voor het schrijven van eenvoudige rapporten en het bijhouden van een log van je werkzaamheden als datascientist zijn zowel gespecialiseerde tekstverwerkers als het schrijven van html in een teksteditor minder geschikt. Beide vragen nogal wat aandacht doordat ze veel muisklikken vragen om tekst op te maken of, in geval van het gebruik van html, je voortdurend moet controleren of je de tags wel juist gebruikt en afgesloten hebt. Doordat ze zo veel aandacht vragen wordt je, zeker bij het loggen, voortdurend van de kern van je werk afgeleid. Gelukkig is er een alternatief dat het mogelijk maakt met minimale inspanning goed leesbare documenten te schrijven. Dit alternatief heet [markdown](#). Markdown is een markup language (net als html), maar omdat [het aantal tags minimaal is en veel eenvoudiger van structuur](#) zijn dan die van html verstoort het gebruik hiervan je denkproces minder.

Door de eenvoudige structuur is markdown heel goed te gebruiken in teksteditoren. Verder zijn er utilities zoals [pandoc](#) waarmee markdown documenten makkelijk om te zetten naar andere formaten zoals html (en daarmee EPUB), ODF of PDF.

Omdat het aantal tags in markdown erg beperkt is loop je al snel tegen beperkingen aan. In standaard markdown moet je bijvoorbeeld html-tags gaan gebruiken als je een tabel wil maken en er is geen mogelijkheid om een index te genereren. Hierom zijn er verschillende markdown dialecten ontstaan die elk hun eigen gebruikersniche ondersteunen. Het dialect waarmee je in dit thema kennis gaat maken, is [rmarkdown](#) van [rstudio.com](#).

Opdrachten

Download en bestudeer de [rmarkdown cheat sheet](#) en de [rmarkdown reference](#). [Duckduck](#) (of Google) naar andere informatiebronnen. Er zijn ook diverse videotutorials te vinden, bijvoorbeeld [deze](#), [deze](#) en [deze](#). Mocht je er, ondanks alle online hulpbronnen, nog steeds vragen hebben, stel ze dan gerust aan je docent.

Opdracht 1: open RStudio en maak een nieuw r markdown document

- waarvan de default output PDF is
- noem dit document '**Opdracht_1.Rmd**'
- Maak een titel van het type **header 3** met als tekst iets origineels zoals 'Tabel'.
- Maak een tabel met minimaal vijf regels en twee kolommen. Zet in de eerste kolom een cijfer (bijvoorbeeld 1) en schrijf in de tweede kolom op dezelfde regel de naam van dat cijfer (bijvoorbeeld één).
- maak een woord vetgedrukt
- maak een woord cursief
- Voeg nog een titel (header 3) toe met als tekst: 'Gedichtje'
- Zet de volgende tekst in je document en maak er een 'block quote' van. N.B. je wil hier RStudio dwingen de regelovergangen op bepaalde plaatsen te houden.

Ik zit me voor het raam
stierlijk te vervelen
ik wou dat ik twee hondjes was
dan kon ik samen spelen!

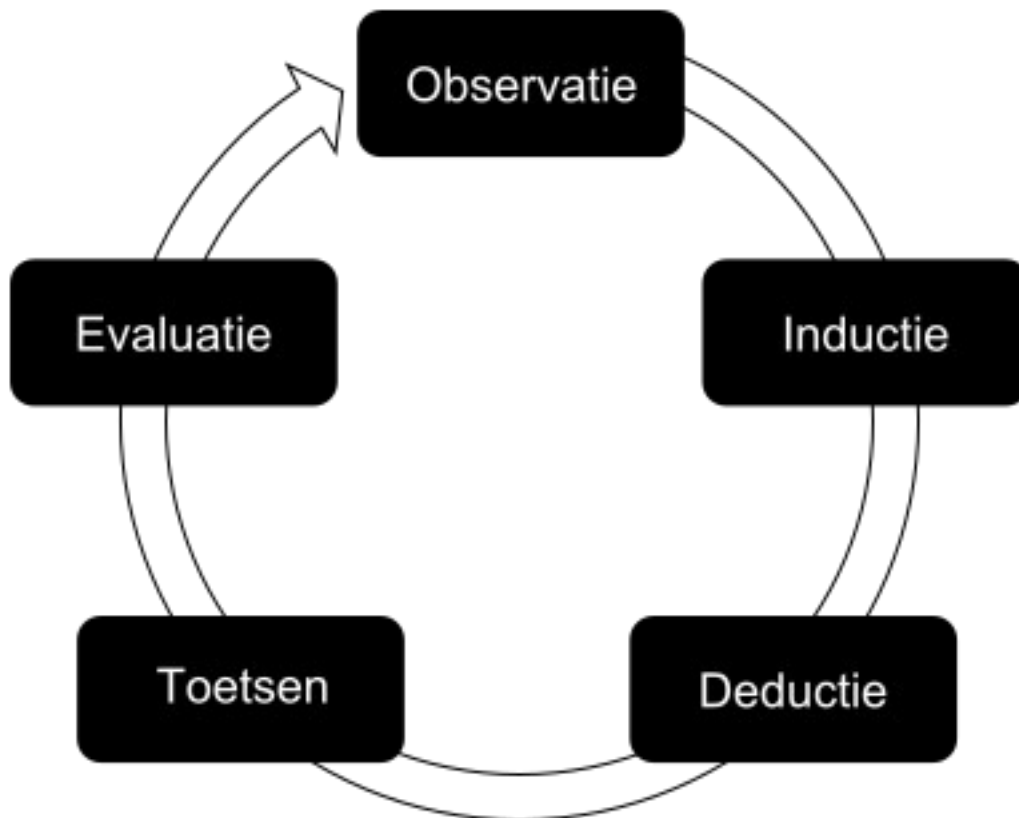
- zoek een plaatje dat vrij is van copyright van een hondje en voeg dit toe aan je document.
- Voeg een inhoudsopgave (toc) toe

Opdracht 2:

- voeg Nederlandse spellingscontrole toe aan je RStudio.
- controleer je document op spelfouten. N.B. Let op! De woordenschat van deze spellingscontrole is beperkt. Accepteer niet zomaar alle suggesties!
- Zoals je ziet zijn de opmaakmogelijkheden RStudio niet geweldig. Je kunt je document ook in LibreOffice Writer openen en daarmee opmaken en je spelling controleren. Tekstverwerkers zoals Writer en Word bieden namelijk uitgebreidere opmaakmogelijkheden dan Markdown.
 - knit je document naar Word formaat
 - open het geëxporteerde bestand in Writer
 - controleer de spelling
 - zoek uit hoe je een bestand uit writer kunt exporteren als .pdf
- nog een alternatief is zelf een Nederlandse *.dic* file te downloaden en ergens in je home te zetten, bijvoorbeeld `~/.config/dict/` en vervolgens RStudio dat bestand te laten gebruiken voor de spellingscontrole. Woordenboeken zijn, bijvoorbeeld, te vinden op [github](#) of [hier](#).

Onderzoek doen

Binnen de (natuur)wetenschappen wordt eigenlijk steeds een cirkel gevolgd, de *empirische cyclus*, zoals die in figuur 1 (met dank aan Wikipedia) staat afgebeeld.



Op basis van *waarnemingen*, die je verkrijgt door *observatie* van de werkelijkheid, wordt door middel van logisch redeneren (*Inductie*) een algemene theorie opgesteld over hoe die werkelijkheid functioneert. Vervolgens wordt er door middel van *deductie* een *hypothese* (een veronderstelling) geformuleerd die op basis van de opgestelde theorie iets voorspelt over de werkelijkheid. Vervolgens wordt deze *hypothese* getoetst door middel van één of meer experimenten. De resultaten van de experimenten worden geëvalueerd (kloppen de uitkomsten met de theorie of niet? Zijn er op basis van de verzamelde gegevens andere theorieën mogelijk die de uitkomsten beter verklaren?) en wordt de originele theorie als nodig aangepast of eventueel zelfs volledig verworpen. Waarna de cyclus zich herhaalt. [zie ook, deze pagina](#) en [deze](#).

Gegevens en Bias

In dit proces is het van het grootste belang dat er zoveel mogelijk geprobeerd wordt om de invloed van eigen vooroordelen c.q. vooringenomenheid te beperken. Een voorbeeld van vooringenomenheid, oftewel *bias*, is dat het meeste medische onderzoek uitgevoerd is/wordt op blanken. Zie bijvoorbeeld de volgende links

<https://link.springer.com/article/10.1007/s00251-016-0914-1>
<https://minorityhealth.hhs.gov/omh/browse.aspx?lvl=4&lvlid=15>
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4354806/>
<https://www.ncbi.nlm.nih.gov/pubmed/?term=Underlying+genetic+structure+impacts+the+association+between+CYP2B6+polymorphisms+and+response+to+efavirenz+and+niraparib#>
<https://www.scientificamerican.com/article/how-to-fight-race-and-gender-bias-in-science-editorial/>

Let wel dat deze bias niet altijd het gevolg hoeft te zijn van latent racisme van de onderzoekers. In de Verenigde Staten van Amerika is bijvoorbeeld gezondheidszorg voor velen niet te betalen en aangezien de blanken zwaar oververtegenwoordigd zijn in de hogere inkomensklassen, zijn ze ook zwaar oververtegenwoordigd in de ziekenhuisbezoeken die de monsters voor veel medisch onderzoek leveren. Mondiaal gezien geldt eigenlijk hetzelfde. Landen met van origine een blanke bevolking zijn vaak rijker en hebben dus meer geld voor onderzoek, waardoor ...

Evaluatie

Behalve dat de waarnemingen (= gegevens, = data) zelf biased kunnen zijn, kan ook de interpretatie van die data niet zuiver gebeuren. In, bijvoorbeeld, de politiek veel gebruikte redeneringen als

'Het kan toch niet zo zijn dat ...'
'Iedereen weet toch dat ...'
'Fake Fact!'
'Dat is iets voor klimaatdrammers!'

hebben in de wetenschap geen plaats omdat ze niet van feiten maar emoties (en dus vooroordelen) uitgaan en worden **drogrede** genoemd.

Observeren en Publiceren

Bij het de stappen 'Observatie' en 'Inductie' van de empirische cyclus (zie figuur 1) wordt in de wetenschap veel gebruikt gemaakt van observaties en ideeën van anderen. Om dit mogelijk te maken moeten

die ideeën en observaties (data) wel beschikbaar zijn. Dit gebeurt door het publiceren van artikelen waarin onderzoek en de uitkomsten daarvan in relevante wetenschappelijke tijdschriften. Voordat een artikel gepubliceerd wordt, wordt het eerst beoordeeld door een andere vakdeskundige of het onderzoek goed uitgevoerd is.

Elk tijdschrift gebruikt z'n eigen layout voor publicaties, maar bevatten (bijna) allemaal dezelfde onderdelen:

- een korte **samenvatting** zodat je niet het hele artikel hoeft te lezen om te bepalen of het voor jou interessant is.
- een **inleiding** waarin uitgelegd wordt welke theorieën/observaties (observatie en inductie in figuur 1) getoetst gaan worden met welke onderzoeksvraag (deductie) en waarom die onderzoeksvraag relevant is.
- om het voor andere onderzoekers mogelijk te maken om de beschreven experimenten en data te controleren op bias c.q. de experimenten te herhalen moet er een **Materiaal en Methoden** sectie aanwezig zijn waarin *en detail* beschreven staat hoe het onderzoek uitgevoerd is en welke datasets hiervoor gebruikt zijn. Dit is van wezenlijk belang omdat hiermee bias in datasets of mogelijke **fraude** boven water gehaald kunnen worden.
- de verkregen resultaten worden zo kaal mogelijk gepresenteerd in de **Resultaten**-sectie
- om vervolgens in de **Discussie** vergeleken te worden met observaties van anderen. Op basis van dit alles wordt beargumenteerd of de onderzoeksvraag beantwoord kan worden en wat dit antwoord dan zou moeten zijn. Ook worden hier op basis van de resultaten suggesties gedaan voor verder onderzoek.
- Het antwoord op de onderzoeksvraag wordt nogmaals samengevat in de **Conclusie**
- Let er op dat alle 'observaties' die je gebruikt in je inleiding of in je discussie voor anderen te controleren moeten kunnen zijn. Dit betekent in de praktijk dat, als je gebruik maakt van werk van anderen je dit duidelijk maakt door een referentie naar de bewuste publicatie. Alle gebruikte referenties worden in de **Literatuurlijst** opgenomen.

De **template** die wij voor verslaglegging gebruiken is een variant van bovenstaand patroon.

Om een publicatie leesbaar te houden moeten alle niet relevante zaken er buiten blijven. Dit kunnen bijvoorbeeld experimenten of resultaten zijn die wel uitgevoerd zijn, maar bij nader inzien toch niet relevant blijken te zijn voor het beantwoorden van de

onderzoeksvraag. Of resultaten gebruikt worden in de discussie is hiervoor de lakmoesproef.

De Onderzoeksvraag

Het beantwoorden van een onderzoeksvraag is dus waar het bij het doen van onderzoek om draait. Niet alle vragen zijn (goede) onderzoeksvragen. Een goede onderzoeksvraag heeft onder andere de volgende kenmerken:

- zij is **onbevooroordeeld** (de vraag stuurt het antwoord niet een bepaalde kant op)
- zij is gericht op het verkrijgen van nieuw inzicht
- zij is te beantwoorden in de beschikbare tijd
- zij is (in principe) beantwoordbaar.

Omdat uitkomsten in de wetenschap vaak lastig te interpreteren zijn, gebruiken we vaak statistiek om ons te helpen om onderzoeksvragen te beantwoorden. Op open vragen ('Hoe...?', 'Waarom...?') kun je echter met statistiek geen antwoord geven. Statistiek kan je wel helpen om uit twee alternatieven de **meest waarschijnlijke** te kiezen. Er wordt bij het statistisch toetsen er van uit gegaan wordt dat je twee elkaar uitsluitende **hypothesen** hebt, de nul-hypothese waarvan je aanneemt dat die waar is en de alternatieve hypothese die je accepteert als je de nul-hypothese verwierpt. Hierom kan het vaak duidelijker zijn om je onderzoeksvraag voor een bepaald experiment als hypothese/alternatieve hypothese paar te verwoorden in plaats van als open vraag. Het op deze manier formuleren van een onderzoeksvraag helpt ook om te bepalen wat voor experiment je uit kunt voeren om haar te beantwoorden en wat voor gegevens hiervoor nodig zijn.

Opdrachten

Verzamel de antwoorden op deze opdrachten in het markdowndocument wat je in opdracht 1 en 2 gemaakt hebt.

Opdracht 3 Een veel gebruikte drogreden is het '[autoriteitsargument](#)'. Beargumenteer of het gebruiken van een referentie die je niet gelezen hebt of niet begrijpt een autoriteitsargument is of niet.

Opdracht 4a Geef voor de volgende uitspraken aan wat voor drogredenen ze bevatten:

- 'Het kan toch niet zo zijn dat ...'
- 'Iedereen weet toch dat ...'
- 'Fake Fact!'

- 'Dat is iets voor klimaatdrammers!'

Opdracht 4b Bestudeer het volgende [artikel](#). Welke drogredenen worden er in dit stuk gebruikt?

Opdracht 5 Leg van elk van de onderstaande vragen uit waarom zij wel of geen goede onderzoeksvraag is.

- Waarom moet het nu net op mijn verjaardag regenen?
- Hou je van mij?
- Zijn honden groter dan katten?

Opdracht 6 Wat is de beste onderzoeksvraag? Leg uit waarom! > 'Zijn alle raven zwart?' > 'Zijn er witte raven?'

Opdracht 7 Formuleer zelf een onderzoeksvraag en leg uit waarom deze volgens jou interessant is.

- Let er op dat zij beantwoordbaar moet zijn binnen de tijd van dit practicum
- Wat zijn de hypothese en de alternatieve hypothese?
- Welke gegevens heb je nodig?

Weekopdracht 1

1. Maak een directory die je volgens onderstaande beschrijving een naam geeft:

groep[nr]-[voornaam]-[achternaam]-opdracht[nummer]

waarbij je:

- [nr] vervangt door de subklas waar je in zit (A of B dus)
- [voornaam] vervangt door je roepnaam
- [achternaam] vervangt door je achternaam
- [nummer] vervangt door het nummer van de weekopdracht, in dit geval dus 1.

Een voorbeeld van een correcte directorynaam is dus:

groepA-arne-poortinga-opdracht1

2. Zet het markdownbestand wat je in de vorige opdrachten gemaakt hebt in deze directory en maak hiervan een .pdf en een .doc (een word bestand) die je ook in deze directory zet.
3. maak een .tar.gz bestand van deze directory (zie voor hulp de man page van tar of gebruik bijvoorbeeld [DuckDuckGo](#))
4. Lever het .tar.gz bestand in via de desbetreffende link op BlackBoard.