**Onderzoeksontwerpen**

Door: Marleen Hasselt | Docent: Maaike van der Vleuten | Vakcode: SOW-MTB2005

***Deze samenvatting is afkomstig uit het collegejaar 2017-2018. Het kan zijn dat sommige onderdelen iets verschillen van de huidige tentamenstof. Let hier op!***

**Cursusomschrijving:** Wetenschappelijke publicaties en sociaal-wetenschappelijke onderzoeksverslagen zijn even divers in kwaliteit als in de onderwerpen die zij behandelen. Soms maken auteurs/onderzoekers verkeerde methodologische keuzen, soms is er een beperking ten aanzien van aanwezige data of van tijd en geld, en soms raakt de gebruikte methodologie van een onderzoek achterhaald door voortschrijdend inzicht. Allemaal aspecten die de betrouwbaarheid en de validiteit van de resultaten kunnen beïnvloeden. Niet alleen als student, maar zeker ook in de loopbaan na de studie is het daarom van belang onderzoeksverslagen of wetenschappelijke publicaties op waarde te kunnen schatten.   
In deze cursus krijgen de studenten handvaten die kunnen helpen bij de methodologische evaluatie van onderzoek aan de hand van specifieke beoordelingscriteria. Daarnaast zullen verschillende onderzoeksontwerpen behandeld worden, zoals longitudinaal onderzoek (panelonderzoek, herhaalde cross-secties, event history analyse), multilevel onderzoek, beleidsonderzoek en experimenteel onderzoek. Bij elk type onderzoek zullen de sterke en zwakke punten belicht worden en hoe je als onderzoeker de valkuilen van elk ontwerp kan ontwijken. Voorbeeldonderzoeken op het terrein van sociologie zullen worden gebruikt om kritisch te leren lezen en te bestuderen hoe onderzoekers om gaan met de zwakheden van hun ontwerp.

**Doelstellingen:**

* De student is in staat overeenkomsten en verschillen aan te geven tussen diverse onderzoeksontwerpen, alsmede de sterke en zwakke punten van de ontwerpen te benoemen.
* De student kan een onderzoeksverslag evalueren op de daarvoor in aanmerking komende methodische aspecten.
* De student kan, gegeven een onderzoeksprobleemstelling, een beargumenteerde keuze maken voor een adequaat onderzoeksontwerp.

**Hoorcollege 1 Introductie en Evalueren van onderzoek**

De volgende onderzoeksdesigns komen aan bod: multilevel, experimenten in surveys, longitudinaal onderzoek  
Tentamen is 70% van je eindcijfer en presentatie 30%.

**Fasen van onderzoek**

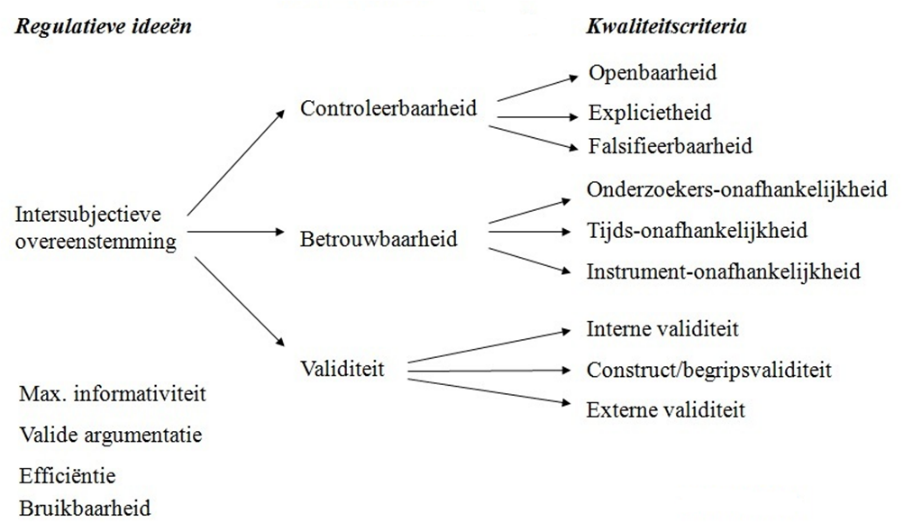
1. Probleemstelling
   1. Wat wil ik weten (vraagstelling)
   2. Waarom wil ik dit weten (doelstelling)
2. Onderzoeksopzet
   1. Keuze onderzoeksstrategie
   2. Uitwerking: waar, wanneer, wie, hoe?
3. Data verzamelen
   1. Bij wie en wat?
   2. Waar, wanneer, hoe, door wie?
4. Analyse gegevens
5. Conclusies
   1. Antwoord op probleemstelling
6. Rapportage en evaluatie van werkwijze en uitkomsten

Deze periode gaan we het vooral hebben over de rapportage en evaluatie van werkwijze en uitkomsten.

**Het evalueren van onderzoek doen we als volgt**

* Beoordeling van kwaliteit van de opzet en uitvoering van de afzonderlijke onderzoeksfasen:
  + Heldere doel/ probleemstelling
  + Analyse correct uitgevoerd?
* Past de samenhang tussen de fasen?
  + Past de onderzoeksstrategie bij de probleemstelling?

Om de kwaliteit van onderzoek te beoordelen zijn er methodologische spelregels en kwaliteitscriteria. Swanborn zegt hierover dat achter de verschillen tussen de onderzoeksdesigns een gemeenschappelijke basis van regulatieve ideeën schuilt.   
Regulatief idee = algemeen leidend principe waaraan wetenschappelijke onderzoek aan moet voldoen.

Regulatieve ideeën en kwaliteitseisen van Swanborn,1996 ------------->

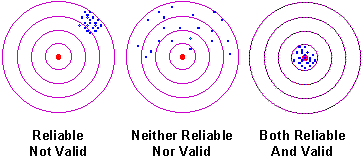
**Intersubjectieve overeenstemming:**   
In hoeverre zijn onderzoekers die hetzelfde stukje van de sociale werkelijkheid bestuderen, het onderling eens? Het gaat hier om een oordeel o.b.v. controleerbaarheid, betrouwbarheid en validiteit.  
Iedere onderzoeker is op zoek naar de waarheid: maar observaties zijn altijd gekleurd door verschillende theoretische invalshoeken, paradigma’s en persoonlijke voorkeuren.   
Het concept van de objectieven en absolute werkelijkheid gebruiken we dus niet, maar we spreken van *voorlopige kennis*.

**Controleerbaarheid:**   
Opzet en resultaten moeten repliceerbaar zijn. Hiervoor moeten de opzet en resultaten: openbaar en expliciet zijn.   
Daarnaast moeten resultaten toetsbaar zijn. Hiervoor moeten uitspraken zowel verifieerbaar (controleerbaar) als falsifieerbaar (je moet het kunnen verwerpen) zijn.

**Betrouwbaarheid:**De mate waarin resultaten vrij zijn vaan niet-systematische (toevallige) fouten.   
De mate waarin herhaalde meting hetzelfde resultaat oplevert. Dit hangt af van:

* Onderzoekers-onafhankelijkheid (codeursbetrouwbaarheid)
* Tijds-onafhankelijkheid (test-retest)
* Instrument-onafhankelijk (parallelle instrumenten)

**Validiteit**De mate waarin resultaten vrij zijn van systematische fouten.  
Meet je wat je wilt meten? Dekt het de werkelijkheid? Je hebt verschillende soorten validiteit:

* Interne validiteit: causale interpretatie van een gevonden samenhang. Zijn alternatieve verklaringen uitgesloten?
  + Inhoudsvaliditeit: het meetinstrument dat je gebruikt bevat alle of ten minste vele soorten items die er volgens deskundigen die je zijn voorgegaan (collega’s betrekken, discussiëren).
  + Soortgenootvaliditeit: als het oude meetinstrument niet meer ‘van deze tijd’ is moet er een nieuw meetinstrument ontworpen. Daarna moet er een soort vooronderzoek gedaan worden waarbij het oude en het nieuwe instrument voorlegt aan een doelpopulatie waarbij je het zelfde item meet. Als er een sterke samenhang optreedt tussen de antwoorden is er sprake van soortgenootvaliditeit.
  + Convergente validiteit: als je als onderzoeker beschikt over twee instrumenten waarmee het zelfde verschijnsel wordt gemeten en ze een sterke samenhang vertonen. Beide instrumenten meten namelijk het zelfde begrip.
  + Construct- of begripsvaliditeit: de mate waarin de empirishe waarnemingen zoals gemeten in gvo en kvo of experiment, de theoretische begrippen/ constructen dekken. Dit is moeilijk te meten. Je kan o.b.v.een onderzoek geen conclusies trekken over de constructvaliditeit van een instrument. Wanneer een instrument in een aantal onderzoeken gebruikt is en de theoretisch verwachte relaties steeds gevonden worden, kan worden geconcludeerd dat de constructvaliditeit kennelijk in orde is.
  + Predictieve validiteit: voor bij psychologische tests/ praktijkgericht onderzoek. Als voorspellingen herhaaldelijk niet uit blijken te kloppen is de predictieve validiteit in het geding.
* Externe validiteit: in hoeverre kunnen de onderzoeksresultaten gegeneraliseerd worden naar de ‘echte wereld’ er zijn twee soorten:
  + Populatievaliditeit: de mate waarin de steekproef een goede afspiegeling is van de populatie waarnaar de onderzoekers willen generaliseren (= representativiteit). Als de steekproef niet representatief is, is de selectie van de onderzoekseenheden niet valide. Aselecte steekproef. Wanneer geen sterke verschillen tussen steekproef en nationale verschillen = representatief.
  + Ecologische validiteit: de mate waarin de resultaten van een onderzoek kunnen worden gegeneraliseerd nar andere plaatsen/ tijden/ omstandigheden als die waarop het onderzoek feitelijk betrekking had. Dit is alleen te meten door het onderzoek echt onder die andere omstandigheden te repliceren.

Onderzoek evalueren a.d.h.v. een voorbeeldstudie: *Gezond werk* door Schalk, Drash, Maas.   
Wanneer we dit onderzoek willen evalueren kijken we naar:

1. Probleemformulering:
   * Arbeidsomstandigheden ------+------> gezondheid
   * Variëteit, uitdaging, salaris ------+------> gezondheid
   * Werklast ------\_------> gezondheid
   * Baanzekerheid ------+------> gezondheid
   * Sociale ondersteuning ------+------> gezondheid
   * Het negatieve effect van werklast op gezondheid is sterker bij een lager salaris
   * Het negatieve effect van werklast op gezondheid is steker bij een kleinere regelruimte.

Commentaar:  
- Centrale onderzoeksvragen niet expliciet uitgeschreven  
- Psychosociale aspecten behelzen maar een deel van het begrip. Fysieke aspecten ontbreken.  
- Het belang van de organisatie wordt niet geproblematiseerd, alleen vaan het beroep/ functie, maar deze wordt uitgeoefend binnen een organisatie.

1. Literatuuroverzicht van eerder onderzoek
2. Theorieën en hypothesen
3. Het onderzoeksdesign
   * Data: time competition survey, cross-sectioneel, geneste structuur, de partners van de werknemers. Kun je causale uitspraken doen?

Commentaar:  
- Arbeidsomstandigheden -> gezondheid of vice versa (causaliteit/ interne validiteit).   
- 30 grote bedrijven (externe validiteit)  
- Vooral kennisintensieve organisatie (externe validiteit)  
- In NL in 2003 (ecologische validiteit)  
- Niets bekend over selectieve non-respons (externe validiteit).

1. De dataverzameling en meetinstrumenten (selectie van eenheden, externe validiteit; methode van data verzameling; meetinstrument)
   * Twee op rationalisatie voor gezondheid:  
     - Aantal dagen ziek in de afgelopen 6 maanden  
     - Hoe goed is uw gezondheid in vergelijking met andere mensen van hun leeftijd?

Commentaar:  
+ twee metingen van hetzelfde begrip  
- twee metingen van hetzelfde kunnen worden gebruikt om een meer betrouwbare schaal te construeren en dat is niet gedaan.   
- definitie van ziekte en de zelfboordeling laten ruimte over voor verschillende interpretaties (construct validiteit).  
- de vergelijking met mensen van eigen leeftijd zorgt er voor dat men in staat moet zijn om in te schatten hoe gezond leeftijdsgenoten zijn (betrouwbaarheid).

1. Data-analyse:
   * Ordinale logistische regressieanalyse:  
     - afhankelijke variabelen zijn ordinaal en op het individuele niveau ingeschat.   
     - Het effect van arbeidsomstandigheden geschat op gezondheid, rekening houdend met enkele achtergrondkenmerken, waaronder opleiding.

Commentaar:  
- Geen rekening gehouden met de bedrjifscontext  
- Controle voor opleiding niet voldoende  
- Causaliteit: is gezondheid wel de afhankelijke variabele of komen ongezonde menen ook vaker in minder aantrekkelijke banen terecht.

1. Resultaten, conclusie, discussie
   * Werklast ------\_------> gezondheid
   * Functiezekerheid ------+------> gezondheid
   * Werkdruk ------\_------> gezondheid. Dit wordt sterker naarmate men meer regelruimte heeft.

Commentaar: zelfde als onder data-analyse.

Algemeen commentaar:  
- Externe validiteit: populatievaliditeit, ecologische validiteit, selectieve respons  
- Interne validiteit: causaliteit, constructvaliditeit

Centrale vragen bij het evalueren van onderzoek:

* Op welke wijze boekt men vooruitgang t.o.v. eerder onderzoek? Theoretisch interessant maar weinig toevoeging wegens methodologische problemen.
* Past het onderzoeksontwerp bij de probleemstelling en wat is de meerwaarde van dit onderzoeksdesign t.o.v. andere designs? Onderzoeksdesign laat geen causale uitspraken toe; beter was geweest en paneldesign te gebruiken.
* Wat zijn de gevolgen van de werkwijzen van de onderzoekers voor de betrouwbaarheid, interne, externe validiteit, bruikbaarheid? Veel methodologische problemen die vooral interne en externe validiteit in de weg staan. Bruikbaarheid is daarom gering.

**Hoorcollege 2 Multilevel Onderzoek**

Multilevel onderzoek is een onderzoeksstrategie die zich richt op de relaties tussen kenmerken van verschillende niveaus (=levels).   
Geneste/ hiërarchische datastructuur: eenheden van lagere niveaus (individuen) delen een hoger niveau (landen):

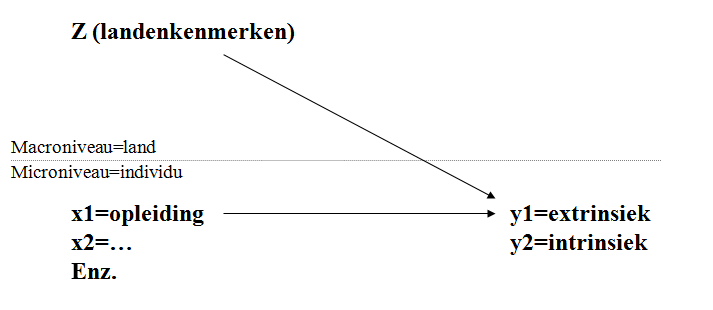
* Leerling – klas – school
* Individu – wijk – stad – regio – land
* Student – opleiding – faculteit – universiteit
* Individu – gezin – familie

Waarom is multilevel onderzoek zo populair?  
- theoretisch relevanter  
- tegenwoordig (sinds de jaren 70) praktisch haalbaar  
- beleidsmatig relevant

Historie:

* Durkheim: de optiek van het individu naast de optiek van de context.
  + Lidmaatschap godsdienstige groepering ------\_------> zelfdoding
  + Wonen in een streek waar bepaalde godsdienstige groepering veel vóórkomt ------\_------> zelfdoding
* Davis: single level analyse. Inactieve deelnemers of inactieve discussiegroep?   
  Actief zijn = individueel kenmerk & zitten in een groep met actieve mensen = context kenmerk.

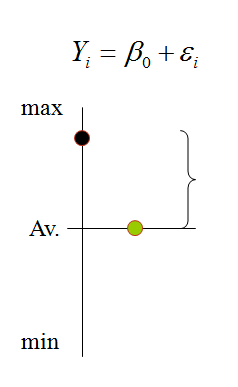
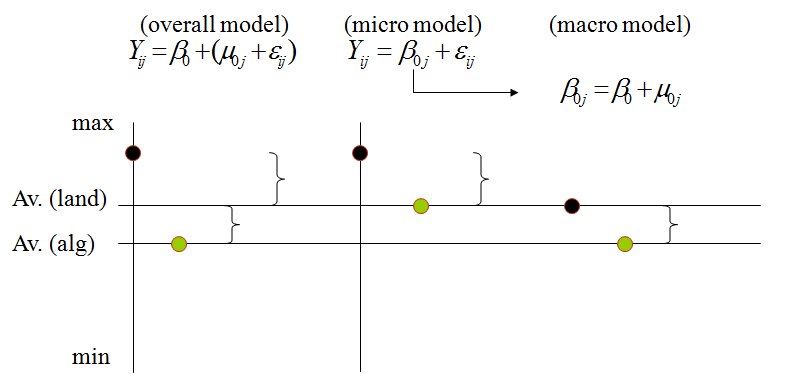
**Voorbeeld artikel:** *Job preferences in Europe* van Gesthuizen en Verbakel  
Data: European Values Study, 1999/ 2002. 22 landen.  
Afhankelijke variabelen:  
- Voorkeur voor extrinsieke baankenmerken: werk dient om doelen buiten het werk te bereiken, zoals materiële doeleinden  
- Voorkeur voor intrinsieke baankenmerken: doel van werk ligt in het werk zelf: persoonlijke ontwikkeling en je talenten benutten.

Conclusie: binnen landen zijn er verschillen tussen individuen in de mate waarin zij belang hechten aan extrinsieke en intrinsieke baankenmerken.   
We willen deze individuele verschillen (variantie binnen groepen) verklaren, maar dan wel door de invloed uit te schakelen van het land waarin een individu woont.

Kortom: *binnen* landen verschillen individuen in de mate waarin zij een voorkeur hebben voor extrinsieke en intrinsieke baankenmerken. Dit wordt ook wel de **variantie binnen groepen** genoemd.  
De gemiddelde mate van de voorkeur voor extrinsieke en intrinsieke baankenmerken verschilt eveneens *tussen* landen. Dit wordt ook wel de **variantie tussen groepen**genoemd.

De verhouding tussen deze variantiecomponenten geeft aan hoe belangrijk de context is.   
**Intraklasse correlatie** = de mate waarin waarnemingen van individuen binnen groepen (hier landen) afhankelijk van elkaar zijn = landvariantie/ (landvariantie + individuele variantie) = (*u*0*j*)/(*u*0*j*+ *e*0*ij*). Met andere woorden: de mate waarin de variantie in voorkeur voor baankenmerken tot stand komt door de invloed van het land.

Hoe ziet dit eruit in een vergelijking?

1. Regressievergelijking zonder hiërarchische structuur en zonder predictoren: single level null model (model 0).   
   Iemands score op extrinsiek of intrinsiek = de gemiddelde score op extrinsiek of intrinsiek, plus zijn of haar afwijking van dat gemiddelde.   
   Het gemiddelde is voor iedereen hetzelfde, de afwijking van het gemiddelde wordt voor iedereen apart geschat. Dit laatste wordt ook wel de error term/ random deel genoemd: datgene dat het model onverklaard laat. Deze individuele afwijking wordt omgezet in een variantiemaat.
2. Regressievergelijking met hiërarchische structuur en zonder predictoren: two-level random intercept null model (model 1).   
   Multilevel: onverklaard deel op individueel en op groepsniveau; de gemiddelde score op de voorkeur voor extrinsieke en intrinsieke baankenmerken varieert over landen.
3. Regressievergelijking met hiërarchische structuur en met predictoren op het individuele niveau: two-level random intercept model (model 2).
   1. Micro-model: 
   2. Macro-model
   3. Overall model: 4

X= kenmerk op het individuele niveau  
Te schatten parameters: B0 B1.  
De variantie-componenten zijn: σ²ε, σ²μ0

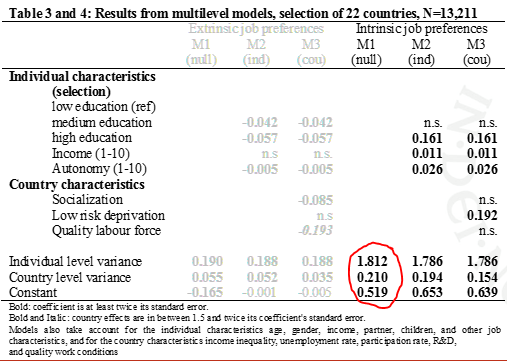
1. Regressievergelijking met hierarchische structuur en met predictoren op het individuele niveau en contextuele niveau: two-level random intercept model (model 3)

micro-model: 

macro-model: 

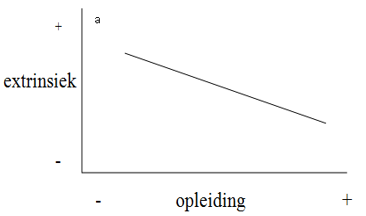
overall model: 

Z=kenmerk groepsniveau  
Te schatten parameters: B0 B1 B2  
De variantie-componenten zijn: σ²ε, σ²μ0

**De resultaten van de multilevel regressiemodellen**

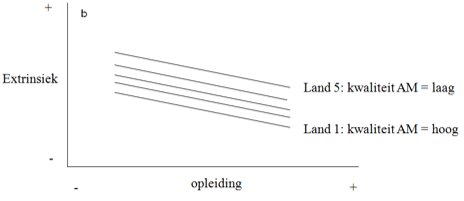
Intraklassecorrelatie: 0,210/(1,812+0,210)= 0,10

Hoeveel van de oorspronkelijke verschillen worden verklaard door dit model?  
(0,210 – 0,154) / 0,210 = 0,27

**Grafische weergave individuele en contextuele effecten**

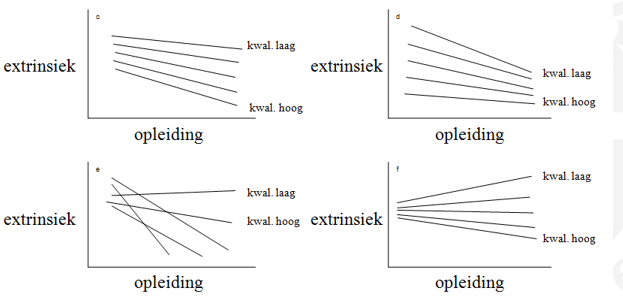
1e niveau: werknemer:  
y = belang extrinsieke baankenmerken  
x = opleidingsniveau.

Zoals je de lijn a nu ziet is een weergave van een single-level model. De lijn zou horizontaal recht lopen als er geen individueel effect zou zijn.

1e niveau: werknemer:  
y = belang extrinsieke baankenmerken  
x = opleidsingsniveau.

2e niveau: land  
z = kwaliteit van de arbeidsmarkt

De lijnen b geven de verschillen weer tussen landen m.b.t. gemiddelde score op de afhankelijke variabele (random intercept model). Als er geen contextueel effect zou zijn dan zou de lijn parallel lopen aan deze lijnen maar tussen land 1 en 5 in.

In c-f: verschillen tussen landen m.b.t. gemiddelde score én het effect van de individuele variabele (hellingshoek varieert per land) (interactie-effecten).

Selectie van eenheden: cluster VS getrapt.

* Clustersteekproef: bij een clustersteekproef ondervraag je een hele groep, meestal reeds bestaande groepen zoals een klas of team
* Getrapte steekproef: uit een clustersteekproef wordt een aselecte groep genomen waarbij je weer een steekproef afneemt.

Bij multilevel onderzoek kan er gebruik worden gemaakt van meerdere type variabelen:

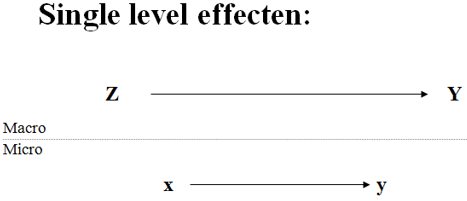
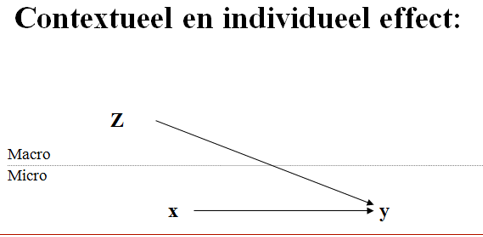
* Globaal kenmerk =Direct gemeten op betreffende niveau, zonder verwijzing naar andere eenheden of niveaus (geslacht of schooltype)
* Analytisch (groeps)kenmerk = Verkregen door aggregatie van globale gegevens van eenheden van een lager niveau (percentage meisjes in een klas)
* Relationeel kenmerk = Relatie van een eenheid met andere eenheden op hetzelfde niveau (mate van contact tussen 2 buren)
* Structureel (groeps)kenmerk = Verkregen door aggregatie van rationele gegevens van eenheden van een lager niveau (aantal contacten tussen bewoners in een wijk)
* Contextueel kenmerk = verkregen door disaggregatie van gegevens van een hoger naar lager niveau (leerling zijn in een grote school)

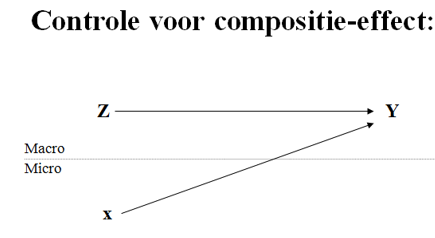
**Single-level vs. Multilevel analyse**

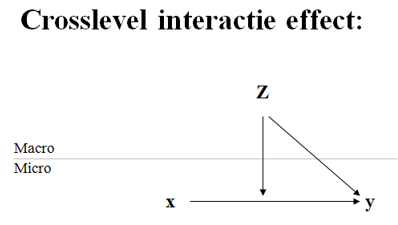
1. Aantal analyse niveau’s
   * Single level/ contexuele analyse: alle analyses op 1 niveau.   
     
   * Multi-level: simulate analyse van meerdere niveau’s  
     
2. Typen variabelen
   * Single level: vaak beperkt tot geaggregeerde groepskenmerken
   * Multi level: alle soorten kenmerken
3. Intraklasse correlatie
   * Single level: geen controle
   * Multi level: wel controle
4. Interpretatie van resultaten
   * Single level: oppassen voor ecologische fout of atomistic fallacy.
     + Ecologische fout: o.b.v. geaggreggeerd niveau conclusies trekken over individuen.
     + Atomistic fallacy: o.b.v. analyse op individueel niveau conclusies trekken over een hoger niveau.
   * Multi level: geen problemen

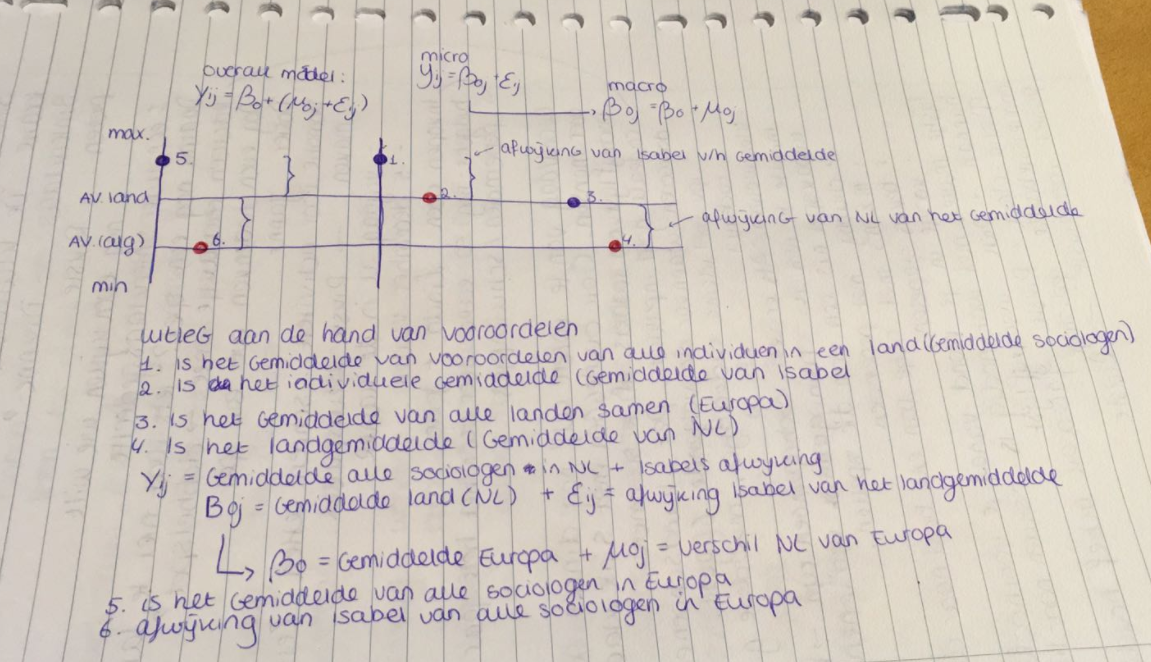
Voorbeeld ecologische fout: als je vindt dat in protestante provincies het zelfmoordcijfer hoger is, concluderen dat protestanten een grotere kans hebben om zelfmoord te plegen dan katholieken.  
Voorbeeld atomistische fout: als je vindt dat in protestanten vaker zelfmoord plegen, concluderen dat in landen met een hoger percentage aan protestanten de gemiddelde kans op zelfmoord hoger ligt dan hoger is.

**Vraagstellingen** (zie ook Snijders en Bosker, 1999)  
x Onafhankelijke vairabele op lager niveau.  
Z Onafhankelijke variabele op hoger niveau.  
y Afhankelijke variabele op lager niveau.  
Y Afhankelijke variabele op hoger niveau.



Het effect van de omgeving is uitgezuiverd voor compositieverschillen tussen groepen.



Extra uitleg van two-level random intercept null model

**Hoorcollege 3 Experimenten en Survey experimenten**

Vandaag:

1. Experimenteel design
   * Pre-experiment
   * Quasi-experiment
   * Zuiver experiment
2. Survey-experimenten

**Doelstelling van experimenteel onderzoek**   
Het bewijzen van een causale relatie tussen te manipuleren onafhankelijke variabelen en te meten afhankelijke variabelen.   
De onderzoeker:  
- Brengt systematisch en gecontroleerd variatie aan in een onafhankelijke variabele;  
- Stelt de effecten van die variatie vast en een andere, afhankelijke variabele;  
- Houdt daarbij alle andere variabelen constant.

Opzet:  
- Wie: onderzoeker verdeelt proefpersonen over groepen.  
- Waar: in het veld, lab of binnen een survey.  
- Hoe: verschillende methodes: observaties, vragenlijsten, etc.

**Het verschil tussen een experiment en een gewone survey**Voorbeeld 1: verband interetnisch contact en vooroordeel.

1. Survey: (mogelijk schijnverband)  
   1. Meting van interetnisch contact  
   2. Meting van vooroordeel
2. Experiment: (causaliteit ligt vast)  
   1. Manipulatie van interetnisch contact  
   2. Meting van vooroordeel

Voorbeeld 2: verband rommelige omgeving en vooroordeel

1. Survey  
   1. Meting van buurt waarin mensen wonen  
   2. Meting van vooroordeel
2. Experiment  
   1. Manipulatie van omgeving  
   2. Meting van vooroordeel

Voorbeeld 3: “Leidt blootstelling aan een voorlichtingsfilm over asielzoekers tot een positievere houding tegenover asielzoekers?”  
Voorwaarden oorzakelijk verband  
- Statistisch verband: X -> Y  
- Chronologische volgorde (film (t0) -> houding tegenover asielzoekers(t1))  
- Uitsluiten schijnverband  
De interne validiteit is hoog als de experimentele variabele vrijwel zeker van invloed is op de afhankelijke variabele.

1. Tonen voorlichtingsfilm (X)
2. Meting houding tegenover asielzoekers (Y)

**Kenmerken van een experiment**

X variabele = experimentele variabele, de enige variabele die systematisch varieert.  
Y variabele = de effect/ gevolg variabele

* Voormeting + nameting
  + Zelfde respondenten
  + Meting voor en na de manipulatie
  + Geen controle groep -> testeffect
  + Groep 1 O1 X O2
* Experimentele + controle groepen
  + Verschillende groepen
  + Experimentele groep is blootgesteld aan manipulatie
  + Controle groep is niet blootgesteld aan de manipulatie
  + Maar geen voormeting
  + *Groep 1* X O1

*Groep 2* O2

Controle voor storende variabelen: uitsluiten alternatieve verklaringen:

* Onafhankelijke variabele varieert systematisch
* Alle andere variabelen worden gelijk gehouden over de controle groep en experimentele groep
* Hoe? -> toewijzingsprocedures
  + Randomisering (R) -> per toeval aan experimentele en controle groep -> storende factoren over de groepen verdeeld.
    - *Groep 1* R X O1  
      *Groep 2* R O2
  + Matchen (minder variantie tussen groepen):
    - Precisie: voor elke proefpersoon in experimentele grep, een vergelijkbare in de controle groep.
    - Globaal: een aantal belangrijke kenmerken gelijkmaken tussen experimentele en controle groep.
  + Homogeniseren: beperken tot een selectie van respondenten (minder variantie binnen en tussen groepen

**Designs van experimenten**

1. Pre-experiment 1
   * Geen randomisering
   * Geen vergelijkingsbasis (of de voormeting of de controlegroep)
   * Één groep + één stimulus, één nameting
   * *Groep 1*  X O1
   * Je weet niet zeker of Y reageert op X
2. Pre-experiment 2
   * Geen randomisering
   * Eén groep, één stimulus, voor- en nameting:
   * *Groep 1* O1 X O2
   * Je weet niet zeker of Y reageert op X
3. Pre-experiment 3
   * Geen randomisering
   * Twee groepen, één stimulus, alleen nameting, geen randomisering:
   * Groep 1 X O1  
      Groep 2 O2
4. Quasi-experiment
   * Geen randomisering
   * Voor praktijkgericht / beleidsonderzoek
   * Sterkere vergelijkingsbasis
   * Twee groepen, één stimulus, voor- en nameting, geen randomisering:
     + Design 1. O1 X O2 experimentele groep   
        O3 O4  controlegroep
   * Eén groep, één stimulus, meerdere voor- en nametingen, geen randomisering:
     + Design 2. O1 O2 O3 O4 O5 X O6 O7 O8 O9 O10
   * Twee groepen, één stimulus, meerdere voor- en nametingen, geen randomisering:
     + Design 3. O1 O2 O3 O4 O5 X O6 O7 O8 O9 O10 experimentele groep

O1 O2 O3 O4 O5 O6 O7 O8 O9 O10 controlegroep

* + Negatief: Testeffect, tijdsintensief
  + Quasi-experimenten zijn vergelijken met zuivere experiment doorgaans:  
    - minder intern valide (veel storende factoren)  
    - meer extern valide (lijken op de echte situaties/ vinden plaats in de echte wereld).

1. Zuiver experiment 1: *Pretest-posttest control-group design* (klassiek experimenteel ontwerp)
   * *Groep 1* R O1  X O2 experimentele groep *Groep 2* R O3 O4  controlegroep
   * MAAR: voormetingen kunnen van invloed zijn op nametingen (nadelig voor interne en externe validiteit).
   * Plafond/ vloer effect (als je al een 10 voor wiskunde hebt, en je krijgt bijles, krijg je geen 11).
2. Zuiver experiment 2: *Posttest-only control group design*
   * *Groep 1* R X O5 experimentele groep

*Groep 2* R O6 controlegroep

1. Zuiver experiment 3: Solomon four group design
   * *Groep 1* R O1  X O2 experimentele groep

*Groep 2* R O3 O4  controlegroep

*Groep 3* R X O5 experimentele groep

*Groep 4* R O6 controlegroep

* + Maar: duur en tijdsintensief

1. Zuiver experiment:
   * *Groep 1* R O1  X1  O2 experimentele groep

*Groep 2* R O3 X2 O4  experimentele groep

*Groep 3* R O5 X3 O6 experimentele groep

*Groep 4* R O7 O8 controlegroep

Interne validiteit is hoog als:  
- de experimentele variabele vrijwel zeker van invloed is op de afhankelijke variabelen;  
- als alternatieve verklaringen uitgesloten kunnen worden.   
Storende factoren spelen vooral een rol bij pre-quasi-experiment. Zuiver experiment vaak hoger interne validiteit. Wel belangrijk is de random toewijzing.

Interne validiteit: storende factoren:  
- Tussentijds extern voorval  
- Rijping: mensen gaan nadenken = kortere tussentijd oplossing  
- Test-effect: oplossing = langere tussentijd of geen voormeting  
- Instrumentatie: oplossing = gelijk houden  
- Selectie: oplossing = randomisatie  
- Selectieve uitval: oplossing = kortere tussentijd.   
Een zuiver experiment controleert voor al deze factoren behalve uitval.

Externe validiteit is hoog wanneer de resultaten gegeneraliseerd kunnen worden naar een populatie en naar andere omstandigheden.   
Een experiment heeft een hoge interne validiteit, lage externe validiteit door dat er zeer selectieve steekproeven worden genomen. Bovendien is er een zeer specifieke stimuli.

Externe validiteit: storende factoren.  
- Interactie voormeting-stimulus (mensen gaan over dingen nadenken)  
- Interactie selectie-stimulus (niet representatieve steekproef)  
- Reactieve locatie (dubbel blind experiment) (mensen reageren op locatie)  
- Belangrijk: controle op storende variabele interne validiteit staan los van de controle op de storende variabelen van de externe validiteit.

Betrouwbaarheid is hoog wanneer:  
- de waarnemingen nauwkeurig en stabiel zijn;  
- waarnemingen vrij zijn van toevallige fouten  
Alle condities zoveel mogelijk gelijk houden

**Survey experiment**

Een experiment binnen een survey noemen we een *survey-experiment.*Definitie: Een systematische manipulatie van een onafhankelijke variabele teneinde een causaal verband vast te stellen. Daarbij worden respondenten *at random* verdeeld over één of meerdere groepen / condities.

Technieken: CAPI, CATI, CAWI, CASI. Met de computer kun je meer mensen bereiken.  
In dit type survey-experimenten krijgen respondenten op basis van toeval verschillende versies van een vraag voorgelegd:

* Experimenten waarin het attitude object wordt gevarieerd, terwijl de context gelijk blijft (experiments that modify choices).
* Experimenten waarin het attitude object gelijk blijft, terwijl de context wordt gevarieerd (experiments that modify the context of a choice). (Piazza & Sniderman, 1998).

Interne validiteit: storende factoren  
- Order en spill-over: bij meerdere experimenten in één survey: effecten van eerdere experimenten vertekenen de resultaten van latere experimenten. Oplossing = random volgorde.  
-Controle manipulatie: is het effect van de manipulatie als verwacht? Kan een out-group benefit ook als ene in-group cost worden opgevat (focus groep)

-Selectieve uitval: minder controle op uitval tijdens een survey dan tijdens een lab-experiment.

Externe validiteit: storende factoren  
- Duur van de effecten  
- Enkelvoudige manipulaties  
- Controlegroep

**Voorbeeld artikel***Does race matter in neighbourhood preferences? Results from a video experiment.*

Onderzoeksvraag: “Heeft de compositie van de wijk (zwart/blank/gemengd) een invloed op wijkvoorkeuren van blanke Amerikanen, onafhankelijk van sociale klasse?”  
Data: Detroit Area Studies, 2004 (N=734) + Chicago Area Studies, 2004-2005 (N=789)  
Methode: CAPI; video-CASI  
Analayse: Multilevel (aparte modellen voor blanke en zwarte Amerikanen).

Afhankelijke variabele: ‘How would you rate the neighbourhood you just saw as a place to live?’  
Onafhankelijke variabele:   
- Video (buurt): Sociale klasse + Etnische samenstelling (gemanipuleerd): blank, zwart of gemixed.  
- Respondent (individu): Negatieve stereotypes van zwarte Amerikanen + Negatieve stereotypes van zwarte wijken + Gevoelsmeting voor de houding ten aanzien van zwarte Amerikanen + Controle variabelen

Goede interne validiteit? Waarom sterk design?

* Overlap klasse en etniciteit: hier variëren ze onafhankelijk
* Random toewijzing van respondenten aan neighbourhood composition (selectie)
* Controle conditie (individual baseline)
* CASI: kleinere interviewer effecten (sociale wenselijkheid)
* Check van de manipulatie (focus groups)!
* Alleen variatie ‘race’; kenmerken en activiteiten constant
* Subtiele manipulatie

Goede externe valaiditeit? Waarom sterk design (volgens de onderzoekers?)

* Potentiële kopers bekijken de wijk op deze manier

Resultaten:

* Wijkvoorkeuren van blanke Amerikanen. Blanke wijken worden het meest positief beoordeeld, gevolgd door gemengde wijken en zwarte wijken.
* Effect ‘race’ sterker voor blanke Amerikanen die een negatieve houding over zwarte Amerikanen hebben.
* Dit experiment houdt het midden tussen een ‘gewoon’ experiment en een ‘survey-experiment’.

**Hoorcollege 4 Longitudinaal onderzoek**

Longitudinaal onderzoek is onderzoek waarbij je op meer dan een tijdstip gegevens verzameld met als doel een verschijnsel door de tijd heen te bestuderen. Je hebt verschillende longitudinale onderzoeken:

1. Analyse van herhaalde cross-sectie (o.a. cohort-analyse)
2. Panel onderzoek
3. Tijdsreeksen

Voordelen longitudinaal onderzoek:

* Veranderingen in kaart brengen
* Betere bewijsvoering met betrekkingen tot causale vraagstukken

**Analyse van herhaalde cross-secties/ cohort-analyse**

Er worden meerdere cross-sectionele studies geanalyseerd. Op meerdere tijdstippen en dwarsdoorsnede van de doelpopulatie, met op elk tijdstip een andere steekproef van onderzoekseenheden. Het resultaat hiervan is een stapeling van cross-sectionele surveys (pooled-dataset).

Nadelen:

* Datapunten zijn niet aaneensluitend maar gespreid
* Geen zicht op individuele veranderingen
* Geen conclusies op individueel niveau mogelijk.

Voordelen:

* Laat trends zien
* De mogelijkheid om cohort-analyse te doen. Hiermee kunnen verklaringen voor veranderingen worden toegeschreven aan: leeftijds-effecten, cohort-effecten, periode-effecten.   
  De veranderingen kunnen gelijktijdig plaatsvinden: met cohort-analyse kun je de invloed ervan vaststellen.

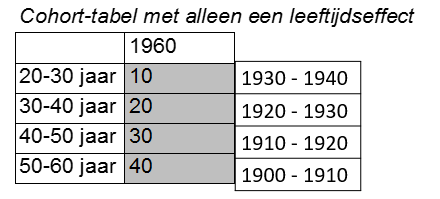
Voorbeeld

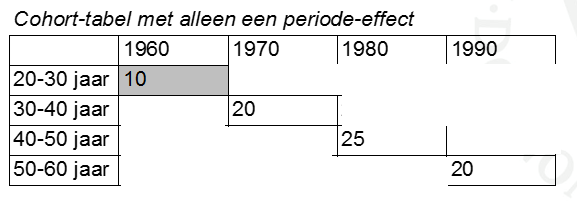
* Begin jaren 90 steunde ongeveer 40% van de Nederlandse bevolking etnische discriminatie.
* Ouderen steunen meer etnische discriminatie

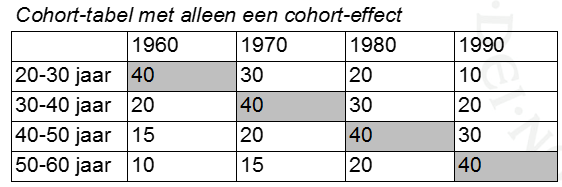
Nu zijn er drie verschillende soorten veranderingen:

1. Leeftijdseffect = naarmate men ouder wordt meer etnische discriminatie?
2. Cohorteffect = oudere generaties meer steun dan jongeren generaties (en blijven dit voor de rest van hun leven doen)?
3. Periode-effect = door historische ontwikkelingen iedereen (jong of oud) meer steun voor etnische discriminatie.

Een speciaal type van cross-sectie analyse is de cohort-analyse. Een cohort is een verzameling mensen die in eenzelfde tijd een belangrijke gebeurtenis meemaken (geboortecohorten, afstudeercohorten).

Leeftijdseffecten: effecten die gepaard gaan met het ouder worden t.g.v. fysieke, psychische en sociale veranderingen die zich fasegewijs afspelen in het proces van ouder worden.

Periode effecten: effecten ie met alle mogelijke gebeurtenissen in de periode tijdens en direct voorde meting te maken hebben en ieders score op de afhankelijke variabele beïnvloeden.

Cohort-effecten: effecten waarbij gebeurtenissen in de tijd cohort-specifiek werken (dus op het ene cohort een andere invloed hebben dan op het andere cohort).

Conclusie

1. Groepen van individuen worden over de tijd heen gevolgd (niet dezelfde mensen)
2. leeftijd-, periode- en cohort effect vereist herhaalde cross-secties
3. Met één cross-sectie kun je ze niet uit elkaar houden.

Schatten van leeftijds-, periode-. en cohorteffecten door middel van inhoudelijke kenmerken.  
Voordelen hiervan zijn:

* Het probleem van perfecte multicollineariteit is opgelost
* Op een inhoudelijke informatieve, theoretische gestuurde wijze
* Die tot toetsbare hypothesen leidt over de invloed van specifieke periode- en cohortkenmerken.

Nadelen hiervan zijn:

* Beschikbaarheid van gegevens
* Vergelijkbaarheid van gegevens over de tijd
* Hebben we alle invloeden van periode/ cohort/ leeftijd kunnen meenemen
* E kenmerken kunnen onderling sterk samenhangen, hetgeen opnieuw voor multi-collineariteit zorgt.

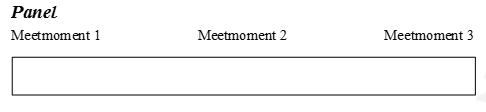
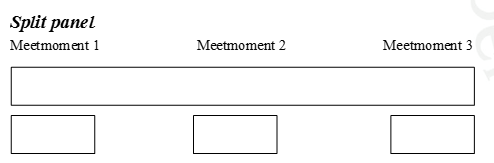
**Panel onderzoek**

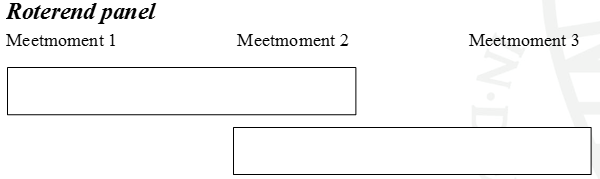
Een panel is een vaste groep personen die bij herhaling voor (dezelfde/ nieuwe) informatie wordt benaderd (ruime/ beperkte zin). Met als doel het vaststellen van veranderingen in opvattingen of gedrag bij een persoon en de invloed van andere factoren daarop. Met andere woorden het vaststellen van individuele veranderingen.

Voordelen van deze methoden: causaliteit vaststellen en individuele veranderingen volgen.

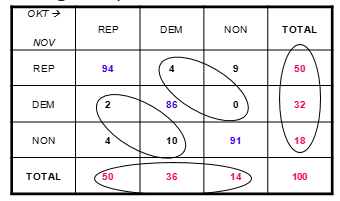
* Paneldesign
  + Grootschalig
  + Hoge mate van structurering van de vragen
  + Kwantitatieve statistische data-analyse
  + Een belangrijke keuze is de keuze van de waarnemingsmomenten en het aantal groepen (wanneer, hoeveel, lengte van de intervallen)
* Panelproblemen
  + Uitval -> betrouwbaarheid, selectiviteit
  + Slijtage van het panel
  + Slijtage van het meetinstrument (migratie achtergrond)
  + Paneleffecten -> positieve trainingseffecten, herinneringseffecten, onbedoelde effecten, vermoeidheidseffecten.
  + Alternatieve verklaringen -> tussentijdse gebeurtenissen, rijping/ groei
* Beperkingen van paneldesign
  + Gering aantal meetmomenten
  + Relatief korte tijdsduur tussen het eerste en het laatste meetmoment.

Er zijn verschillende soorten panel onderzoek, met name in het aantal groepen in het paneldesign:

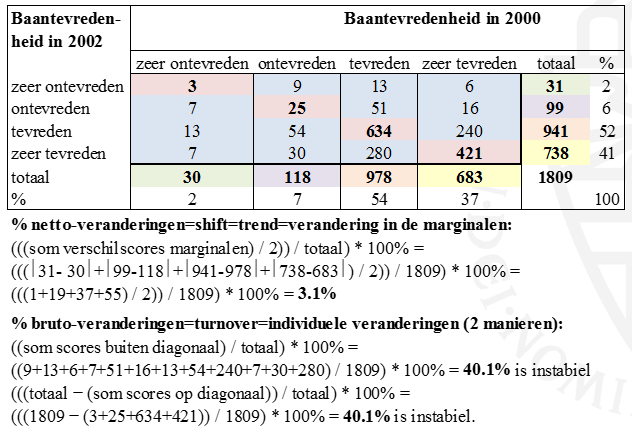




De analyse van panel data:

1. De turn-over tabel:
   1. Structurele of % netto-veranderingen vaststellen: shift (marginalen; horizontaal & verticaal).
   2. Individuele of % bruto-verandering vaststellen: turnover (alles buiten diagonaal).
2. Complexe analyses (meer meetmomenten of meer variabelen): gaan we hier niet op in.

Voorbeeld Turnover Tabel

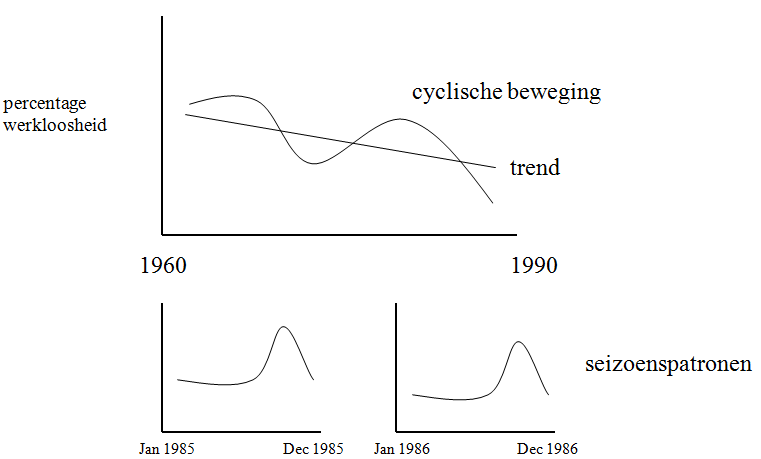


**Tijdreeks-analyse**

Definitie: Een naar de tijd geordende verzameling van *geaggregeerde* waarnemingen over hetzelfde verschijnsel in *verschillende* steekproeven met doorgaans *constante* tijdsintervallen tussen de waarnemingen. N.B. elke verzameling van in de loop der tijd herhaalde observaties zou je in brede zin een tijdreeks kunnen noemen.

Bijvoorbeeld: maandelijkse inflatie cijfers, wekelijkse werkloosheidspercentages, dagelijkse kijkdichtheidscijfer, jaarlijkse krantoplages.

Een tijdreeks bestaat uit cijfers op geaggregeerd niveau; de probleemstelling is dus gericht op veranderingen in geaggregeerde gegevens. Opeenvolgende waarnemingen zijn gerelateerd aan elkaar = autocorrelatie/ seriële correlatie.

Doelstellingen van tijdreeks-analyse:  
Beschrijvingen van de interne structuur. (trend; cyclische ontwikkelingen; seizoenspatronen)

1. Voorspellen van toekomstig gedrag
2. Relatie tussen meerdere tijdreeksen
3. Evaluatie van interventies (interrupted, tijdreeksanalyse/ interventie-analyse).

Aandachtspunten bij analyse van tijdreeksen:

* Groot aantal waarnemingen nodig (>50)
* Lengte van (en regelmatigheid in) de tijdsintervallen;
* Beperkingen van gebruik van secundaire data (mbt. validiteit)
* Geaggregeerde gegevens, dus kans op ecologische fout  
  Ecologische fout = op basis van gegevens over collectiviteiten (bijv. geaggregeerde gegevens) uitspraken doen over eenheden op een lager niveau (bijv. individuen).

Op basis van samenhang tussen geaggregeerde gegevens mag je geen conclusies trekken over individuen.

Voorbeeld van een ecologische fout: hoe hoger het percentage allochtonen in een gebied, des te hoger het percentage PVV-stemmers in dat gebied. Je mag nu niet concluderen dat allochtonen vaker PVV stemmen.