

# Rubric Eindopdracht DATA-PyV NS-109B 2022-2023

Voorbereiding: Py (50 pt)					
Vraag	Ontbreekt (0%)	Onvoldoende (40%)	Voldoende (65%)	Goed (85%)	Uitstekend (100%)
<b>2(a) Lees het locatie bestand in</b> [4 pt]	- data wordt niet ingelezen	- data wordt ingelezen, maar in onwerkbaar format zoals niet gesplitst bij kolommen, of als str/int/nan waar dit niet gewenst is	- lon en lat en afstand tot epi ingelezen als float	zie voldoende + - lijst van afgekorte namen als str ingelezen	zie goed + - compact (niet omslachtig) - goed leesbaar / makkelijk aan te passen (bv geen heel file-path)
<b>2(b) Maak plot</b> [3 pt]	- geen plot gemaakt	- plot heeft minimale lay-out omvat niet alle 3 types data	- plot bevat alle 3 types data - berekening naar decimalen voor Groningen fout - geen/raar gebruik gemaakt van cartesische coördinaten	zie voldoende + - correcte as-labels legenda	zie goed + - figuur word opgeslagen - legenda op handige plek lay-out goed gekozen (helder verschil tussen types)
<b>2(c) Functie die lat lon geeft vanuit station code</b> [4 pt]	- geen coördinaten gevonden	- geen functie geschreven coördinaten gevonden door op te zoeken	- functie geschreven juiste coördinaten komen uit functie - <u>niet</u> volgens voorbeeld (bv input is geen str en/of output zijn geen twee losse floats	- functie geschreven - juiste coördinaten komen uit functie - volgens voorbeeld	zie goed + - compact (niet omslachtig) of - robust (checkt data-types; geeft zinnvolle errors)
<b>3(a) Functie om station-data in te lezen</b> [4 pt]	- geen data ingelezen	- geen functie geschreven data kan wel worden ingelezen	- functie geschreven waarmee data ingelezen wordt - input is <u>geen</u> 4-letter str - output is <u>niet</u> één np.array	- functie geschreven waarmee data ingelezen wordt - input is 4-letter str - output is één np.array	zie goed + - compact (niet omslachtig) of robust (checkt data-types; geeft zinnvolle errors)
<b>3(b) Lees data in</b> [3 pt]	- data niet ingelezen	- data ingelezen d.m.v. kopiëren van veel dezelfde code (uit a)	- eigen functie van a gebruikt om data in te lezen op een rare manier (bv functie gekopieerd i.p.v. alleen aanroepen)	- eigen functie van a gebruikt om data in te lezen	zie goed + - alle 3 stations in één keer ingelezen: bv met lijst van str van station namen en loop
<b>3(c) Bereken baseline kanaal 1 BAPP</b> [3 pt]	- geen baseline berekent	- gemiddelde berekent over verkeerde range geen onzekerheid berekent of onzekerheid geschat/helemaal verkeerd berekent	- gemiddelde berekent over verkeerde range - onzekerheid berekent over zelfde verkeerde range - SDOM gebruikt i.p.v. stdev	- gemiddelde en onzekerheid correct berekent - over juiste range	zie goed + - variabelen (range, naam, kanaal) zijn gemakkelijk aan te passen (bv functie geschreven, of variabelen gedefinieerd)
<b>3(d) Verschuif data voor baseline</b> [3 pt]	- data niet verschoven	- (sommige) data verschoven met <u>een ander getal</u> dan berekent bij c - niet alle data verschoven (bv alleen kanaal 1, of alleen BAPP, of alleen eerste 20 sec)	- alle data verschoven met de baseline uit c - baseline niet aangepast per kanaal en/of per station	- alle data verschoven met de bijbehorende baseline	zie goed + - functie geschreven en correct gebruikt - alle data in één keer ingelezen: bv met lijst van str van station namen en loop
<b>3(e) Pot 3 sub figuren</b> [4 pt]	- geen figuur geplot of - 3 losse figuren geplot met elk een andere data zonder opmaak	- 3 losse figuren geplot met elk een ander kanaal van BAPP of - 1 figuur geplot met 3 data-sets over elkaar; data is niet te onderscheiden figuren hebben aslabels	- 1 figuur geplot met 3 data-sets over elkaar - figuren hebben juiste opmaak (correcte aslabels, legenda; data sets zijn te onderscheiden)	- 1 figuur geplot met 3 subplots met gedeelde x-as - figuur heeft correcte aslabels, legenda/titels per sub figuur	- figuur is mooi vorm gegeven, bv: grid op achtergrond, gelijk y-assen, ... - figuur wordt opgeslagen
<b>3(f) Schrijf functie om t_s te vinden</b> [5 pt]	- geen t_s gevonden - t_s geschat uit ruwe data (geen functie, niet geautomatiseerd)	- gevonden t_s is fout: 1 sigma gebruikt, opdracht verkeerd begrepen, fout door gebruik van ruwe data ipv baseline-verschoven data, ...	- juiste t_s kan gevonden worden - geen functie, wel gemakkelijk in gebruik door bovenaan variabele namen aan te passen (zoals data_lijs, tijd_lijs, aantal_sigma, tijd_sigma, ...) of - functie neemt en/of geeft andere I/O dan gevraagd	- juiste t_s kan gevonden worden met functie die als input minimaal 2 lijsten heeft (data en tijd) en als output één t_s geeft	- functie geschreven bij c gebruikt of nu hiervoor een losse functie gemaakt - (kw)args toegevoegd voor #sigma, tijd_sigma -functionaliteit toegevoegd voor bv multi process,robustheid

## Voorbereiding: Py (50 pt)

Vraag	Ontbreekt (0%)	Onvoldoende (40%)	Voldoende (65%)	Goed (85%)	Uitstekend (100%)
<b>3(g) Schrijf functie om onzekerheid in <math>t_s</math> te vinden</b> [5 pt]	- geen onzekerheid in $t_s$ gevonden	- (vrijwel) handmatig i.c.m. werk uit f waarden in onzekerheid van $t_s$ gevonden	- geautomatiseerd i.c.m. werk uit f onzekerheid van BAPP correct gereproduceerd - geen functie geschreven	- functie geschreven die voor een gegeven station (input) de correcte waarde van $t_s$ als output geeft	zie goed + - compact (niet omslachtig) of - robust (checkt data-types; geeft zinvolle errors)
<b>3(h) Plot <math>t_s</math></b> [4 pt] Let op: wélke data x-as en y-as zijn is niet gegeven, en afstand tot epi op x-as is dus niet fout	- geen plot gemaakt	- plot gemaakt van handmatig getypte data (met onzekerheden, hand getypt of 0.1 sec) - geen as-labels foute eenheden	- geautomatiseerd data verkregen voor plot, maar met veel herhaling waar een loop mogelijk was - gebruik gemaakt van 0.1 sec onzekerheid in $t_s$ - correcte as-labels	- handig gebruik van loop voor verzamelen van data voor de plot - gebruik gemaakt van $t_s$ uit f (juist of onjuist) - correcte as-labels	zie goed + - handig gebruik gemaakt van alle eerder gemaakte functies - figuur wordt opgeslagen
<b>3(i) Fit <math>t_s</math> - epi</b> [3 pt]	- geen fit gedaan		- verkeerde functie gefit (niet lineair, of zonder y-as-afsnede) of - verkeerd toegepast (bv onzekerheden niet meegenomen)		- correct lineaire fit gedaan
<b>Kwaliteit van het script</b> [5 pt]	- lastig te volgen (bv geen comments) - (veel) syntax fouten	- minimaal commentaar - variabelen hebben onhandige namen (te kort of veel te lang)	- combinatie van comments en variabele namen maakt het voldoende leesbaar	- eigen functies worden correct gebruikt waar nodig - commentaar helder en bondig - oogt overzichtelijk	- code is efficiënt zonder leesbaarheid te verminderen - comments geven extra inzicht en/of beschrijven extra functionaliteit goed

## Live-deel: V (50 pt)

Vraag	Ontbreekt (0%)	Onvoldoende (40%)	Voldoende (65%)	Goed (85%)	Uitstekend (100%)
<b>4.2(a) lat lon van 5 stations</b> [3 pt]	Locaties niet gegeven	Slechts output in script	Informatie wordt in antwoorddocument geven, maar is niet goed formateerd	Informatie wordt in antwoorddocument geven en is wel goed formateerd (bv tabel)	Zo als bij goed, maar het wordt bij data naar figuur verwezen of bij figuur naar data (goed gebruik van figuur en data)
<b>4.2(b) Highlight 5 stations</b> [3 pt]	Plot niet gedaan	Slechts output in script	Bijna goed opgemaakte figuur, maar meer dan 1 elementen/details zijn niet helemaal goed of slechts locaties van 5 stations zijn geplotted ipv alle stations en met herkenbaar marker	Goed opgemaakte figuur: - legende en goed te onderscheidend markers - bijschriften (as, figuur) duidelijk en goed leesbaar	Zo als bij goed, maar het wordt bij data na figuur verwezen of bij figuur na data (goed gebruik van figuur en data)  onderschrift meer dan minimaal (volledige zinnen, een beetje context)
<b>4.3(a) Rapporteer baselines</b> [5 pt]	Geen rapportage of compleet foute rapportage (eenheid ontbreekt, afronding meer dan 2 stations fout)	Slechts output in script of meer dan 3 fouten in rapportage of goede rapportage maar van data voorbereidingsdeel	Bijna correcte rapportage, maar meer dan 1, minder dan 3 fouten of rapportage with standaard deviatie ipv SDOM of baseline correct, maar voor 20 ipv 10 seconden	Correcte rapportage van alle 5 stations  1 (reken) fout toegestaan	Correcte rapportage van alle 5 stations zonder fout; data goed geformateerd
<b>4.3(b) Beschrijf SDOM en gemiddelde</b> [6 pt]	Geen antwoord of antwoord compleet fout	Definities niet correct	Alleen correcte definities gegeven, maar geen of foutieve beschrijving	Definities correct, beschrijving gemiddelde + SDOM voor rapportage	Zo als bij goed, maar meer dan minimaal uitleg (volledige zinnen, context)
<b>4.3(c) Uitleg 5 sigma</b> [6 pt]	Geen antwoord of antwoord compleet fout	2 sigma geeft 5% (1:20) significantie maar geen verband met aantal data punten	2 sigma geeft 5% (1:20) significantie; er zijn veel datapunten, dus zou dit tot een 'toevallig' beginmoment leiden. (te algemeen verhaal)	2 sigma geeft 5% (1:20) significantie; er zijn 20000 (of order van grootte) datapunten, dus zou dit tot een 'toevallig' beginmoment leiden	Zo als goed, maar 5 sigma (3e-7, 1:3.5 miljoen) wordt uitgerekend (bv met cdf) in aan aantal datapunten gerelateerd.
<b>4.3(d) Rapporteer t_s</b> [5 pt]	Geen rapportage of compleet foute rapportage (eenheid ontbreekt en afronding meer dan 2 stations fout)	Slechts output in script of meer dan 3 fouten in rapportage of goede rapportage maar van data voorbereidingsdeel	Bijna correcte rapportage, maar meer dan 1, minder dan 3 fouten of rapportage with standaard deviatie ipv SDOM of baseline correct, maar voor 20 ipv 10 seconden	Correcte rapportage van alle 5 stations  1 (reken) fout toegestaan	Correcte rapportage van alle 5 stations zonder fout; data goed geformateerd
<b>4.3(e) Rapporteer snelheid</b> [6 pt]	Geen rapportage of compleet foute rapportage (eenheid ontbreekt en afronding fout)	Slechts output in script (plot + rapportage)	Correcte rapportage met externe onzekerheid (zie commentaar)		Correcte rapportage met interne onzekerheid (zie commentaar)
<b>4.3(f) interne- of externe covariantiematrix</b> [5 pt]				Correcte definitie & antwoord interne onzekerheid	Zo als goed, maar goed geredeneerde argumentatie
<b>4.3(g) Leg uit: correlatie</b> [5 pt]	Geen antwoord		Alleen waarde gerapporteerd	Definitie van correlatie (bij fit) en waarde gerapporteerd	Zo als bij goed, maar opgemerkt dat correlatie hier niets betekend omdat het een lineair fit met weinig datapunten is
<b>4.3(h) Leg uit: <math>\chi^2</math></b> [6 pt]	Geen antwoord	Definitie $\chi_{red}^2$	Definitie $\chi_{red}^2$ ; moet rond de 1 liggen	Definitie $\chi_{red}^2$ ; moet rond de 1 liggen, $\chi^2$ verdeling wordt genoemd	Definitie $\chi_{red}^2$ ; moet rond de 1 liggen, $\chi^2$ verdeling wordt genoemd en juist beschreven, significante afwijking afhankelijk van degrees of freedom

\* als een plot niet zo als gevraagd in het antwoord document staat, maar anders goed gemaakt is in de script: -1 trede (maar niet minder dan voldoende)

NB: interne onzekerheid vereist weinig data punten (check) maar consistent  $\chi^2$  ( $\sim 1$ ). Als een student het  $\chi^2$  verdeling toepast en een significante afwijking van 1 vaststelt, dan mag ook de externe onzekerheid gebruikt worden. (groepen A en F)