

# Fouten in Excel

Marco Steen

Excel is een mooi programma en er kan bijna alles mee. Dat betekent niet dat er geen problemen mee kunnen zijn. Een beetje complex spreadsheet heeft net zo veel mogelijkheden om fouten te bevatten als er functies zijn in Excel. Waarschijnlijk nog wel meer.

Naast fouten in formules komen ook verkeerde of onverwachte uitkomsten voor, die het gevolg zijn van de werking van Excel zelf. De uitkomst is dan wel 'as designed' maar niet als gewent.

Tenslotte is het niet ondenkbaar dat er echte fouten in Excel zitten waardoor de uitkomst niet klopt. Voor zo ver ik weet zijn er op dit moment geen voorbeelden van dit soort fouten bekend, maar in Excel 2007 zat ooit wel een fout waardoor in een heel specifieke situatie een verkeerde uitkomst werd gepresenteerd. Het week ook niet een klein beetje af: als de uitkomst 65.535 zou moeten zijn, werd 100.000 weergegeven.

De belangrijkste bron van fouten in Excel, en ieder ander spreadsheetprogramma, is de eigenaar of maker van het bestand. Veel bestanden zijn organisch tot stand gekomen: het begint met een relatief eenvoudige tabel met een enkele formule, maar in de loop van de tijd wordt dit aangevuld met steeds meer 'handige' toevoegingen. Waar het er in een eenvoudige tabel niet zo toe doet of een formule een beetje netjes is ingericht, wordt het na verloop van tijd onoverzichtelijk. Dat hoeft niet direct tot fouten te leiden, maar het risico wordt er wel groter van. Voorbeelden van formules met een groot risico op verkeerde uitkomsten zijn:

- Getallen opnemen in formules: **=A1+500**  
In een ad-hoc spreadsheet kan dat handig lijken, maar het is dodelijk als het bestand wordt hergebruikt.
- Vermenigvuldigen met een vaste waarde: **=A1\*1,21**  
Als de minister van financiën bedenkt dat het btw-percentage moet wijzigen, moet je op zoek naar alle cellen waar dat zo is verwerkt.
- Verwijzing naar individuele cellen in plaats van een reeks: **=A1+B1+C1** in plaats van **=SOM(A1:C1)**  
Als er een kolom wordt toegevoegd tussen B en C wordt die 'vergeten' in de telling.
- Een tik- of 'aanklik'-fout bij het maken van een formule: **=A1+B2** waar het **=A1+B1** zou moeten zijn.
- Eén formule aanpassen en vergeten die te kopiëren naar de rest in dezelfde kolom.  
Bijvoorbeeld als je de situatie uit het vorige voorbeeld probeert te repareren.
- Gebruik van complexe formules: **=HYPERLINK('[&DEEL(CEL('bestandsnaam');VIND.SPEC('[;CEL('bestandsnaam'))+1;VIND.SPEC(');CEL('bestandsnaam'))-VIND.SPEC('[;CEL('bestandsnaam'))-1]&']&'&'kredietoverzicht'&'!b8';'naar kredietoverzicht')**

Deze heb ik niet voor dit artikel verzonnen. Hij komt zo uit een bestand dat ik regelmatig gebruik.

Er zijn een paar dingen die je kan doen om verkeerde uitkomsten te voorkomen of te herstellen.

Als eerste een open deur: voorkom dat er fouten in sluipen door gewoon slim te werken. Maak spreadsheets niet complexer dan je zelf kan overzien. Denk niet: 'Dat kan ook zo wel even' om daarna te vergeten wat je hebt aangepast. Een tweede tip om te voorkomen dat er fouten ontstaan, is het gebruik van tabellen in Excel. Door een reeks op te maken als tabel kan je niet alleen eenvoudig de opmaak van de hele tabel aanpassen, maar Excel dwingt ook af dat formules

in een kolom binnen een tabel altijd gelijk zijn. Als je één cel aanpast wordt dat automatisch toegepast op de rest van de kolom.

Order-id	Werknemer	Klant-id	Orderdatum	Verzendingdatum	Verzendingnaam
30	Anne Helling-Larsen	27	15-1-2006	22-1-2006	Verzendbedrijf B
31	Jan Kotas	4	20-1-2006	22-1-2006	Verzendbedrijf A
32	Mariya Sergienko	12	22-1-2006	22-1-2006	Verzendbedrijf B
33	Michael Nappper	8	30-1-2006	31-1-2006	Verzendbedrijf C
34	Anne Helling-Larsen	4	6-2-2006	7-2-2006	Verzendbedrijf C
35	Jan Kotas	29	10-2-2006	12-2-2006	Verzendbedrijf B
36	Mariya Sergienko	3	23-2-2006	25-2-2006	Verzendbedrijf B
37	Laura Giussani	6	6-3-2006	9-3-2006	Verzendbedrijf B

Als je eenmaal een spreadsheet hebt, is het slim om formules te controleren voor je conclusies trekt op basis van de uitkomsten. Gelukkig bevat Excel al een aantal controlefuncties waarmee je een niet al te groot bestand kunt beoordelen. Deze functies zijn te vinden op het tabblad Formules. De eenvoudigste, maar heel handige functies, zijn het aanwijzen van bron- en doelcellen van een specifieke cel. Daarmee kun je snel zien of een formule is gebaseerd op de cellen die je verwacht.

Rijlabels	Defect	Vraag	Wijziging	Eindtotaal	Defect/totaal
België	37	11	5	53	70%
Duitsland	53	16	5	74	72%
Frankrijk	73	14	16	87	99%
Nederland	80	23	24	115	70%
Spanje	40	11	12	63	63%

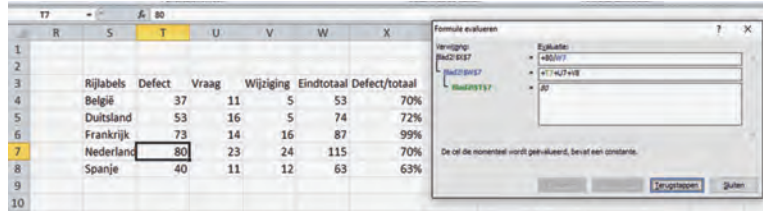
Je selecteert een cel, klikt op 'Broncellen aanwijzen' en Excel geeft aan welke cellen gebruikt worden in de formule. Als je dat voor een aantal cellen doet, zie je snel of er iets mis is.

Rijlabels	Defect	Vraag	Wijziging	Eindtotaal	Defect/totaal
België	37	11	5	53	70%
Duitsland	53	16	5	74	72%
Frankrijk	73	14	16	87	99%
Nederland	80	23	24	115	70%
Spanje	40	11	12	63	63%

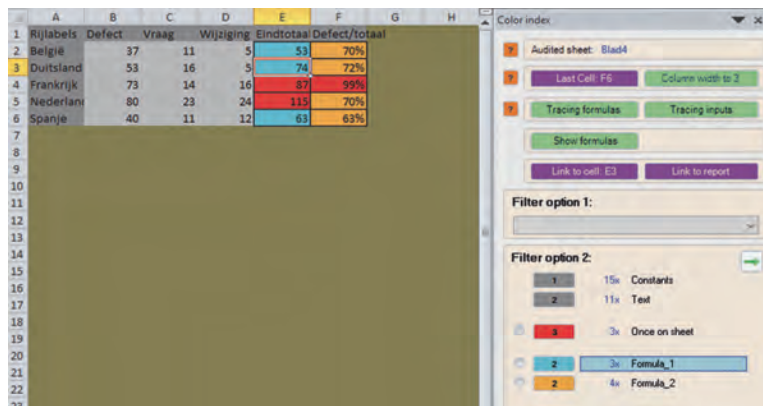
Een stap verder kom je, door de formules weer te geven in plaats van de uitkomst. Je kan dan snel van alle formules in het spreadsheet zien hoe ze zijn opgebouwd. Dit is wat minder intuïtief dan de vorige functie, maar geeft wel meer gedetailleerde informatie.

Rijlabels	Defect	Vraag	Wijziging	Eindtotaal	Defect/totaal
België	37	11	5	=+T4+U4+V4	=+T4/W4
Duitsland	53	16	5	=+T5+U5+V5	=+T5/W5
Frankrijk	73	14	16	=+T6+U6	=+T6/W6
Nederland	80	23	24	=+T7+U7+V7	=+T7/W7
Spanje	40	11	12	=+T8+U8+V8	=+T8/W8

Een derde functie om vooral reeksen naar elkaar verwijzen-de formules te beoordelen vind je onder de knop 'Formule evalueren'. Daarmee kan je stap voor stap door de formules lopen tot de broncellen.



Als je aan de ingebouwde functies van Excel niet voldoende hebt, zijn er speciale programma's beschikbaar om je spreadsheet te beoordelen. Een daarvan is ExcelAnalyser. Dit is een invoegtoepassing die vooral gericht is op de zakelijke markt. Wie wil zien wat de mogelijkheden zijn: er is een volledig werkende probeerversie die 15 dagen gratis gebruikt kan worden. ExcelAnalyser laat met kleuren zien welke cellen een afwijkende formule bevatten. Dat zijn in mijn voorbeeld inderdaad de formules waar iets mee mis is. Verder geeft ExcelAnalyser nog veel meer informatie over de formules, waaronder formules waar getallen in voorkomen.



### Foute uitkomsten door Excel zelf

Eigenlijk is dit de aanleiding voor dit artikel. Ergens vorig jaar kreeg ik via René Suiker de vraag hoe het kan dat berekeningen die logischerwijze exact op nul moeten uitkomen dat in Excel niet altijd doen. De vraag was aangevuld met voorbeelden waarbij in Excel en Quattro Pro onverklaarbare verschillen ontstonden, terwijl in OpenOffice de berekening wel netjes op nul uitkwam. Nu heb ik geen Quattro Pro beschikbaar maar ik kan bevestigen dat dit voorbeeld in Excel en, in mijn geval, LibreOffice tot een verschillende uitkomst leidt, waarbij LibreOffice precies nul als uitkomst heeft en Excel een verschil laat zien. Het is weliswaar een erg klein verschil (iets meer dan 0,0000000000002), maar toch ... Voor wie het voorbeeld wil reproduceren: hierbij de getallen en formules die zijn gebruikt:

A	B
=AFRONDEN(B1;4)	2802,336
=AFRONDEN(B2;4)	1829,9295
=AFRONDEN(B3;4)	=A1-A2
=((A1-A2)-A3)	=((A1-A2)-B3)-0

Als je de eerste twee getallen van elkaar aftrekt houd je het derde getal over. Trek je van dat resultaat het derde getal weer af kom je op nul uit. Ik heb me door mensen die nog kunnen rekenen zonder spreadsheet laten vertellen dat het echt klopt. Alleen blijkt in dit voorbeeld dat Excel ergens een klein stukje overhoudt.

A	B
2802,336	2802,336
1829,9295	1829,9295
972,4065	972,4065
-2,27374E-13	0

Is hier nu iets vreselijk mis met Excel? Nee, dat valt wel mee. Wat er aan de hand is heeft te maken met het feit dat Excel een computerprogramma is en computerprogramma's werken eigenlijk allemaal met een binair (tweetalig) getalstelsel, terwijl wij over het algemeen hebben leren rekenen met het tientalig, decimaal stelsel. Probleem is dat sommige, voor ons heel normale, getallen als 0,1 (of 1/10) niet exact omgezet kunnen worden naar binair. In binair formaat wordt 0,1 benaderd door 0001100110011100110011. Er ontstaan op deze manier altijd afrondingseffecten. Daarbij hanteert Excel getallen met een precisie van 15 decimalen. Dat lijkt (en is over het algemeen) ruim voldoende maar soms levert het dus vreemde afwijkingen op.

### Oplossing

Het is eigenlijk haast nooit nodig om rekening te houden met deze afrondingsverschillen. Voor wie het van levensbelang is om een enorm hoge nauwkeurigheid te hanteren, bijvoorbeeld als je een raket naar een andere planeet wilt sturen, zou ik geen normale spreadsheet aanraden. Ik verwacht dat de meesten onder ons goed kunnen leven met de beperkingen die Excel op dit punt heeft. Het valt eigenlijk alleen op als Excel het nodig vindt om over te schakelen naar de zgn wetenschappelijke notatie, die er uitziet als bijvoorbeeld 1,23E-13. Vaak is het voldoende als je de getalsopmaak aanpast, zodat Excel niet meer dan een of twee decimalen laat zien. Alleen als je verder rekent met de uitkomst van een berekening en het daarbij belangrijk is dat de uitkomst echt gelijk is aan een andere waarde, moet je iets met de formule doen. Dit is bijvoorbeeld het geval als je een tekst wilt weergeven als de waarde '0' is bereikt (=ALS(A4=0;'De koek is op';')). In dat geval moet je in de formule de uitkomst van de berekening eerst afronden op een kleiner aantal decimalen.

### Waarom LibreOffice niet?

Er blijven bij mij twee vragen over waar ik geen antwoord op weet. De eerste is waarom LibreOffice in het voorbeeld uit de vraag wel netjes op 0 uit komt. Daar heb ik eigenlijk nog geen antwoord op kunnen vinden. Ik zie wel dat Excel na 15 decimalen het getal afkapt terwijl LibreOffice het getal afrondt: 5,6 wordt in Excel 5,0 en in LibreOffice 6,0. Het lijkt er op dat LibreOffice een iets andere interpretatie van de standaard hanteert. Dit betekent niet dat het beschreven effect in LibreOffice niet voorkomt. De principiële oorzaak zit ook in LibreOffice ingebakken. De tweede vind ik intrigerender. We hebben gezien dat de formule  $=(A1-A2)-A3$  resulteert in een heel klein getal, maar net geen nul. Halen we de haakjes weg, wat eigenlijk geen verschil zou mogen maken, komt het resultaat wel op nul uit. Dus  $((A1-A2)-A3) <> 0$  maar  $+A1-A2-A3=0$ . Nog mooier wordt het als je bij die laatste formule nul optelt. Dan heb je plotseling weer een waarde:  $+A1-A2-A3+0 <> 0$ . Wie het weet mag het zeggen en, liefst ook begrijpelijk, opschrijven in de volgende SoftwareBus.