

SAS IN ONDERZOEK

BASISBEGINSELEN EN PRAKTIJKTIPS



# SAS in onderzoek

## Basisbeginselen en praktijktips

*Dimitri Mortelmans*

Acco Leuven / Den Haag

*Eerste druk: 2009*

*Tweede, herziene druk: 2016*

*Gepubliceerd door*

Uitgeverij Acco, Blijde Inkomststraat 22, 3000 Leuven, België

E-mail: [uitgeverij@acco.be](mailto:uitgeverij@acco.be) – Website: [www.uitgeverijacco.be](http://www.uitgeverijacco.be)

*Voor Nederland:*

Acco Nederland, Westvlietweg 67 F, 2495 AA Den Haag, Nederland

E-mail: [info@uitgeverijacco.nl](mailto:info@uitgeverijacco.nl) – Website: [www.uitgeverijacco.nl](http://www.uitgeverijacco.nl)

*Omslagontwerp:* [www.frisco-ontwerpbureau.be](http://www.frisco-ontwerpbureau.be)

© 2016 by Acco (Academische Coöperatieve Vennootschap cvba), Leuven (België)

Niets uit deze uitgave mag worden veelevoudigd en/of openbaar gemaakt door middel van druk, fotokopie, microfilm of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever.

No part of this book may be reproduced in any form, by mimeograph, film or any other means without permission in writing from the publisher.



# Inhoud

<b>1. Introductie</b>	13
<b>2. Voorbereiding 1: Twee versies van SAS</b>	15
2.1 SAS University Edition (UE)	15
2.1.1 Virtualbox downloaden en installeren	15
2.1.2 SAS University Edition downloaden en installeren	16
2.1.2.1 Downloaden	16
2.1.2.2 Importeren en configureren in VirtualBox	18
2.1.2.3 SAS University Edition opstarten en gebruiken	21
2.1.3 De interface	23
2.1.3.1 De navigatieruimte	24
2.1.3.2 De werkruimte	25
2.1.4 SAS UE afsluiten	25
2.2 SAS Desktop	27
2.2.1 Het programma installeren	27
2.2.2 De interface	27
2.2.2.1 De navigatieruimte	28
2.2.2.2 De werkruimte	32
2.2.3 SAS Desktop afsluiten	32
<b>3. Voorbereiding 2: De European Social Survey</b>	33
3.1 Toegang tot de ESS	33
3.2 De ESS downloaden	34
3.3 De ESS gebruiksklaar maken voor SAS UE	35
3.4 De ESS gebruiksklaar maken voor SAS Desktop	39
<b>4. Deel 1: Een kennismaking met SAS</b>	43
4.1 Inleiding	43

4.2	Een eerste sessie in SAS UE	44
4.3	Een eerste sessie in SAS Desktop	45
4.4	De SAS-vensters	47
4.4.1	Venster 1: De Editor (Code)	48
4.4.2	Venster 2: Het LOG-venster	50
4.4.3	Venster 3: het OUTPUT-venster – de Results Viewer	51
4.5	Output van SAS lay-outen	52
4.6	Het SAS-huishouden: libraries en databestanden	53
4.7	SAS naar je hand zetten: Opties	56
4.7.1	Opties om de output in het OUTPUT-venster aan te passen	56
4.7.2	Opties om commentaar in het LOG-venster aan te passen	57
4.7.3	Opties groeperen in autoexec.sas	58
4.8	SAS syntax	59
4.8.1	Waarom Syntax gebruiken?	60
4.8.2	Basisstructuur van een programma in SAS	61
4.8.2.1	Algemene principes	61
4.8.3	Programma's voor databeheer	62
4.8.4	Programma's voor statistische procedures	63
4.8.5	Uitzondering: éénregelprogramma's	64
4.9	Moeilijkheden bij het programmeren	64
4.9.1	Veelvoorkomende typfouten	65
4.9.1.1	Puntkomma vergeten	65
4.9.1.2	Aanhalingstekens vergeten	65
4.9.1.3	Verkeerd databestand gebruikt	66
4.9.2	Detecteerbare fouten en problemen	66
4.9.2.1	Een procedure blijft lopen	66
4.9.2.2	Er verschijnt geen output	67
4.9.2.3	Een variabele die aangemaakt werd, is leeg	67
4.9.2.4	Ernstige fouten	68
4.9.3	Syntax structureren en becommentariëren	68
4.9.4	Syntax openen en opslaan	70
4.9.5	Hoe programma's uitvoeren?	70
4.10	Gratis: 1 miljoen SAS-consultants	71
<b>5.</b>	<b>Deel 2: Basisbewerkingen</b>	<b>75</b>
5.1	Het invoeren van externe gegevens	75
5.1.1	De ESS invoeren in SAS	76

5.1.2	De procedure PROC IMPORT	82
5.1.2.1	Data importeren, gescheiden door komma's: CSV	84
5.1.2.2	Data importeren, gescheiden door tabs: TAB	87
5.1.2.3	Data importeren, gescheiden door willekeurig teken: DLM	88
5.1.2.4	Data importeren uit Excel	89
5.1.2.5	Data importeren uit SPSS en STATA	91
5.1.2.6	Data importeren uit R	92
5.1.3	Gegevens importeren met een DATA-stap	93
5.2	Een vragenlijst omzetten in een SAS-databestand	95
5.2.1	Stap 1: Een codeboek aanmaken	95
5.2.1.1	Typen vragen herkennen	96
5.2.1.2	Variabelen namen geven	99
5.2.1.3	Categoriewaarden en missende waarden	102
5.2.1.4	De vorm van het codeboek	103
5.2.1.5	Automatisch een codeboek aanmaken in SAS	105
5.2.2	Stap 2: De variabelen in een SAS-bestand plaatsen	106
5.2.2.1	Het nieuwe bestand aanmaken	106
5.2.2.2	Variabelen toevoegen in het VIEWTABLE-venster	107
5.2.2.3	Variabelen toevoegen met syntax	109
5.3	Variabelen analyseklaar maken	111
5.3.1	Nieuwe variabelen aanmaken en een lengte bepalen	112
5.3.2	Labels hechten aan een bestaande variabele	113
5.3.3	Data leesbaar maken: formats	114
5.3.3.1	Het onderscheid tussen formats aanmaken en toewijzen	115
5.3.3.2	Het probleem van de library 'LIBRARY'	116
5.3.3.3	Het aanmaken van formats	117
5.3.3.4	Het gebruiken/toewijzen van formats	119
5.3.3.5	Ranges definiëren in formats	121
5.3.3.6	Formats gebruiken bij alfanumerieke variabelen	123
5.3.3.7	Lay-out meegeven aan formats	123
5.3.3.8	Permanente toewijzingen van formats verwijderen uit databestanden	126
5.3.3.9	Permanente toegewezen formats niet tonen in procedures	127
5.3.4	Missing Values	128
5.3.4.1	Missings definiëren met formats	130
5.3.4.2	Missings definiëren met databeheer	130
5.3.4.3	Geneste formats maken	131
5.3.5	Een voorbeeld	132
5.4	SAS-gegevens uitvoeren	135
5.4.1	De procedure PROC EXPORT	136
5.4.1.1	Data exporteren, gescheiden door willekeurig teken: DLM	136

5.4.1.2	Data exporteren naar Excel	137
5.4.1.3	Data exporteren naar SPSS en STATA	138
5.4.1.4	Data exporteren naar R	139
5.4.2	Gegevens exporteren met een DATA-stap	140
5.4.3	Uitvoer van frequenties	141
5.4.4	Uitvoer naar een correlatiematrix	143
5.5	SAS Output exporteren	144
5.5.1	Een volledige output wegschrijven in PDF of RTF (SAS UE)	145
5.5.2	Een volledige output wegschrijven in PDF of RTF	146
5.5.3	ODS-tabellen kiezen uit een procedure	148
5.5.4	Output wegschrijven in Excel	149
5.5.5	Output wegschrijven in PowerPoint	153
5.5.6	Output wegschrijven naar een SAS-databestand	158
<b>6.</b>	<b>Deel 3. Databeheer</b>	<b>161</b>
6.1	Inleiding	161
6.2	Werken met bestanden	162
6.2.1	De plaats van databestanden in SAS	162
6.2.2	Nieuwe databestanden maken op basis van bestaande	165
6.2.3	Selecties maken in databestanden	166
6.2.3.1	Het selecteren van variabelen in een databestand	167
6.2.3.2	Het selecteren van cases in een databestand met WHERE	168
6.2.3.3	Het selecteren van cases in een databestand met IF	170
6.2.4	Bestanden samenvoegen	171
6.2.4.1	Het concateneren van bestanden	173
6.2.4.2	Het mergen van databestanden	173
6.2.4.3	Het mergen van bestanden met ingebouwde controle	175
6.2.5	Een bestand kantelen (WIDE en LONG)	178
6.2.5.1	Van WIDE naar LONG	179
6.2.5.2	Van LONG naar WIDE	182
6.2.6	Een bestand aggregeren	184
6.2.7	Een databestand opsplitsen in een analyse	191
6.2.7.1	Het BY-commando	191
6.2.7.2	Het CLASS-commando	193
6.2.8	Een bestand verwijderen of hernoemen	195
6.3	Werken met variabelen	196
6.3.1	Berekeningen op variabelen uitvoeren	196
6.3.1.1	Eenvoudige berekeningen: constanten en operatoren	196
6.3.1.2	Functies gebruiken in berekeningen	199



6.3.1.3	Toepassing 1: Het omkeren van een variabele	203
6.3.1.4	Toepassing 2: Variabelen dichotomiseren	205
6.3.2	Voorwaardelijk databeheer	207
6.3.2.1	Voorwaarden stellen met een SELECT-groep	207
6.3.2.2	Voorwaarden stellen met IF-THEN	209
6.3.3	Gegroepeerde code uitvoeren na een voorwaarde (DO-groep)	210
6.3.4	Overbodige code vermijden: lussen en arrays	210
6.3.4.1	Lussen: DO-groepen	211
6.3.4.2	ARRAYs	213
6.3.4.3	DO-groepen en ARRAYs combineren	215
6.3.4.4	Meerdere ARRAYs in een DO-groep gebruiken	216
6.3.5	Categoriale variabelen hercoderen	217
6.3.5.1	Herocoderen met FORMATS	217
6.3.5.2	Herocoderen met voorwaarden	218
6.3.6	Continue variabelen hercoderen	220
6.3.6.1	Verdeling volgens gelijke klassenbreedte	221
6.3.6.2	Verdeling volgens de nested means methode	223
6.3.6.3	Klassenverdeling op theoretische basis	225
6.3.6.4	Fouten bij klassenverdelingen	225
6.3.6.5	Continue variabelen centreren en standaardiseren	226
6.3.7	Tekstvariabelen hercoderen	228
6.3.7.1	Methode 1: Tekstvariabelen omzetten naar numerieke variabelen	229
6.3.7.2	Methode 2: Herwerken met behulp van Excel	232
6.3.8	Variabelen hernoemen	237
6.4	Werken met cases	239
6.4.1	Cases zichtbaar maken	239
6.4.2	Data sorteren	242
6.4.3	Cases selecteren	245
6.4.4	Steekproeven trekken uit een databestand	246
6.4.5	Cases wege	248
6.5	Toepassing van databeheer: Datacleaning	249
6.5.1	Cleaning op zicht	249
6.5.2	Inhoudelijke cleaning	251
6.5.2.1	STAP 1: De fouten selecteren en zichtbaar maken	251
6.5.2.2	STAP 2: Fout verbeteren	252
<b>7.</b>	<b>Deel 4: Statistische analyses</b>	<b>255</b>
7.1	Tabellenanalyse	255
7.1.1	Frequentieverdelingen	256

7.1.1.1	Een eenvoudige frequentietabel	256
7.1.1.2	Grafieken toevoegen	257
7.1.2	Kruistabellen	258
7.1.2.1	Kruistabellen maken en vormgeven	258
7.1.2.2	Percentages in kruistabellen	261
7.1.3	Grafieken toevoegen	264
7.2	Univariate kengetallen en bivariate associatiematen	265
7.2.1	Kengetallen voor een variabele opvragen	265
7.2.1.1	De UNIVARIATE procedure	266
7.2.1.2	Grafieken toevoegen aan UNIVARIATE	268
7.2.1.3	De MEANS procedure	269
7.2.2	De normaliteit van een metrische variabele nagaan	271
7.2.3	Samenhang van variabelen nagaan	275
7.2.3.1	Associatiematen bij kruistabellen (Chi-kwadraattoets)	276
7.2.3.2	Correlaties berekenen	283
7.2.3.3	Het vergelijken van twee gemiddelden (T-toets)	288
7.2.3.4	Het vergelijken van meer dan twee gemiddelden (ANOVA)	292
7.3	Multivariate analyses	297
7.3.1	Lineaire regressie	297
7.3.2	Logistische regressie	300
7.3.3	Factoranalyse	304
7.4	Analyses uitvoeren in R (met PROC IML)	310
7.4.1	SAS klaarmaken voor R-analyses	310
7.4.2	R-code uitvoeren vanuit SAS	313
<b>8.</b>	<b>Grafieken</b>	<b>317</b>
8.1	Inleiding: een overzicht van grafische mogelijkheden in SAS	317
8.1.1	Grafieken opvragen in procedures	317
8.1.2	Grafieken maken met grafische procedures	318
8.1.3	Menugestuurd grafieken maken in SAS UE	319
8.2	Een overzicht van grafieken	322
8.2.1	Staafdiagram	323
8.2.2	Lijngrafiek	325
8.2.3	Histogram	326
8.2.4	Taart-diagram	327
8.2.5	Box-plot	328
8.2.6	Scatterplot	329
8.2.7	Scattermatrix	331

8.3 Grafische opties bij de grafieken	333
8.3.1 Algemene grafiek-componenten (titels, legendes)	334
8.3.2 Opmaak van staven	335
8.3.3 Opmaak van lijnen	337
8.3.4 Opmaak van assen	340
8.4 Grafieken combineren	340
8.5 Grafieken bewerken in ODS Graphics Editor (SAS desktop)	343
8.6 Grafieken opslaan	345
<b>9. Appendix: ODS-namen van tabellen in SAS procedures</b>	<b>347</b>
9.1 PROC ANOVA	347
9.2 PROC CORR	349
9.3 PROC FACTOR	350
9.4 PROC FREQ	352
9.5 PROC LOGISTIC	356
9.6 PROC REG	359
9.7 PROC TTEST	363
9.8 PROC UNIVARIATE	364
<b>10. Index</b>	<b>365</b>



I.

# Introductie

SAS is een beetje een buitenbeetje geworden. Het programma behoort internationaal tot de meest krachtige statistische verwerkingsprogramma's en is toch slechts beperkt gekend in de Lage Landen. Aan de ene kant krijgt het programma concurrentie van de gebruiksvriendelijkheid van *SPSS*. Aan de andere kant komen nieuwe pakketten opzetten zoals *STATA* of het open source programma *R*. Wat maakt de positie van SAS dan toch zo bijzonder in dit concurrentiële landschap? SAS behoort tot de top van de wereld als het aankomt op databeheer en gevorderde statistische analyses. Het combineert dus de kracht van de nieuwste analyseroutines met een degelijke basis in data management. Die combinatie maakt het programma dermate uniek dat een (her)ontdekking van SAS voor vele onderzoekers een aangename verrassing zal zijn. Daar bovenop heeft SAS internationaal een van de meest uitgebreide gebruikersgemeenschappen. Dat maakt dat nieuwe onderzoekers heel snel en eenvoudig toegang hebben tot een schat aan informatie om zowel eenvoudige als uitermate complexe problemen aan te pakken. De relatief toegankelijke syntaxtaal is hierbij zeker een troef.

Het handboek SAS vertrekt vanuit de meest recente versie van SAS, met aandacht voor de syntaxtaal, vanuit basisbeginselen van statistiek en databeheer. De statistiek noch de programma-interface staan alleen centraal in dit boek. We willen in dit handboek een breed spectrum bieden van taken en problemen waar een onderzoeker voor komt te staan. Uiteraard is kennismaken met een nieuw programma een van deze uitdagingen. Om die reden staan we ook uitgebreid stil bij de interface van SAS en de manier om op een vlotte manier met het programma om te gaan. We reiken daarbij twee versies aan van het programma: de gratis versie "SAS University Edition" en de betalende versie "SAS 9.4". Van daaruit richten we ons dan op de praktijk van het onderzoek zelf. Vertrekkend van het invoeren van gegevens, kijken we naar de basiskenmerken van een goed databestand. We besteden uitgebreid aandacht aan de manier waarop variabelen in een databestand analyseklaar worden gemaakt. Een goede analyse kan immers niet zonder variabelen die degelijk gedefinieerd zijn. Ook databeheer krijgt vanuit die filosofie veel aandacht. We gaan uitgebreider in op databeheer dan een ander handboek statistiek zou doen. We zijn er immers van overtuigd dat onderzoekers al te vaak het bos ingestuurd worden, zonder veel zicht op het bewerken en klaarmaken van

variabelen voor analyses. Databeheer is een cruciale stap in de kwaliteitsbewaking van analyses. Enkel met de juiste basisvariabelen kan je tot correcte conclusies komen. Analyseren zelf komt uiteraard ook aan bod in dit boek, al moesten we hier strenge keuzes maken. De omvang van een handboek SAS laat niet toe om tegelijk ook nog de hele univariate en multivariate statistiek op te nemen. Daarom werd geopteerd om de basistechnieken in dit boek op te nemen: tabellenanalyse, beschrijvende statistiek en enkele multivariate technieken. We staan tot slot ook stil bij de grafische mogelijkheden van SAS in een hoofdstuk over grafieken.

We hopen dat de studenten, onderzoekers en andere lezers na het doornemen van dit boek zonder angst met SAS aan de slag gaan en met voldoende bagage eenvoudige en complexe statistische analyses zullen uitvoeren.

Hoewel alle fouten in dit boek volledig op mijn conto te schrijven zijn, wil ik heel uitdrukkelijk Elke Claessens bedanken voor het nalezen en becommentariëren van eerdere versies van dit boek. Vanuit de outback van Australië kwam een onophoudelijke stroom suggesties en verbeteringen. Heel veel dank is haar deel.

DIMITRI MORTELMANS

2.

## Voorbereiding I: Twee versies van SAS

SAS bestaat in verschillende versies. In dit boek maken we gebruik van twee versies: de SAS University Edition (UE) en de desktop versie van SAS onder Windows (versie 9.4). In dit eerste hoofdstuk gaan we dieper in op de installatie van deze versies en op de interface van beide programma's. De syntax van SAS is uniform over alle versies heen maar er zijn kleine verschillen in het gebruik van databestanden en de interface.

### 2.1 SAS University Edition (UE)

SAS UE is een gratis versie van SAS. In deze versie zijn niet alle analytische modules opgenomen die de desktop versie van SAS rijk is, maar het manipuleren van databestanden en het uitvoeren van de meest voorkomende univariate en multivariate analysetechnieken zijn mogelijk. De versie is gratis omdat SAS academische gebruikers (studenten, docenten en onderzoekers) zo breed mogelijk toegang wil geven tot de software. In deze paragraaf geven we aan waar je het programma kan downloaden en hoe de installatie verloopt.

#### 2.1.1 Virtualbox downloaden en installeren

SAS UE is een zogenaamde “virtual application” (vApp). Dat wil zeggen dat het programma niet als zelfstandig programma wordt geïnstalleerd. SAS UE wordt uitgevoerd *binnen* Oracle VirtualBox. VirtualBox is een zogenaamd virtualisatieprogramma. Dat betekent dat het programma toelaat om andere besturingssystemen te draaien in een afgeschermd omgeving. Dit soort programma's is gemaakt voor software ontwikkelaars. Die kunnen dan in een aparte omgeving een nieuwe versie testen van bijvoorbeeld Windows en kijken of hun programma's in die omgeving nog zullen werken. SAS UE werkt onder Linux. Omdat heel weinig computers het besturingssysteem Linux hebben, werd bijgevolg gekozen om SAS UE in een virtuele machine aan te bieden. Op die manier kunnen gebruikers van Windows of Macintosh toch de software draaien,

zonder apart Linux te moeten installeren. Je moet ook niets van Linux kennen om met SAS UE te werken want het programma wordt via een webbrowser opgestart (zie verder) en je komt niet in aanraking met het besturingssysteem zelf.

Het installeren van VirtualBox is bijgevolg een noodzaak om te kunnen werken met SAS UE. Je kan VirtualBox gratis downloaden op volgende URL:

<https://www.virtualbox.org/wiki/Downloads>

Je dubbelklikt op het installatiebestand en volgt de installatiewizard. Als de installatie klaar is, dan verschijnt volgend scherm:



We laten het programma nu even voor wat het is en concentreren ons in tweede instantie op het downloaden en installeren van SAS UE.

## 2.1.2 SAS University Edition downloaden en installeren

SAS UE is geen zelfstandig programma maar een vApp (Virtuele Applicatie). Dat wil zeggen dat de installatie moet gebeuren *binnen* VirtualBox. Vooraf moet je echter SAS UE downloaden. We beschrijven de verschillende stappen in de volgende paragrafen.

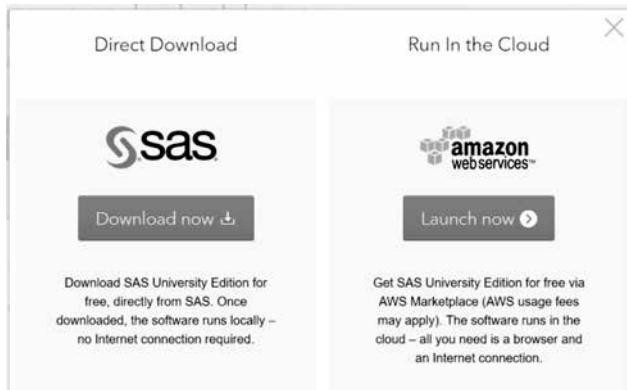
### 2.1.2.1 Downloaden

Om het programma te downloaden, bezoek je volgende URL:

[http://www.sas.com/en\\_us/software/university-edition.html](http://www.sas.com/en_us/software/university-edition.html)

Je klikt op de knop “get free software” en je wordt gevraagd of je de software wil downloaden dan wel uitvoeren in de Cloud. Kies om te downloaden:






Je komt nu terecht op de download-pagina van SAS UE. In verschillende stappen leidt men je doorheen het proces. Je zal merken dat ook het installeren van VirtualBox één van de noodzakelijke stappen is die op de website wordt weergegeven.

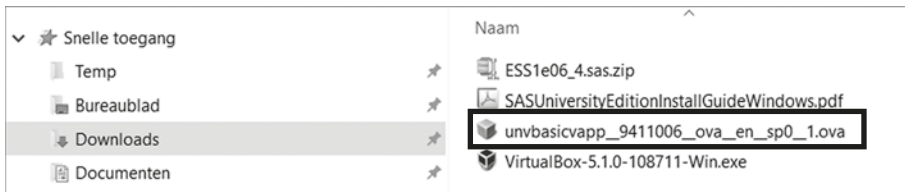
De download van SAS UE vereist zelf drie stappen: (1) je moet een account aanmaken, (2) je moet akkoord gaan met de gebruiksvoorwaarden en (3) je moet het bestand zelf downloaden:



Wanneer je deze stappen doorloopt, zal je zien dat je eigenlijk een virtuele aankoop doet van SAS UE waarvoor je dan 0 \$ betaalt. De registratie brengt je immers naar de SAS Webshop en van daaruit kan je het programma dan downloaden.

	Description	Unit Price	Unit	Total
	<b>SAS University Edition</b> VirtualBox Single user Non-commercial use	0.00 USD	1	0.00 USD
<a href="#">Download</a>				
				<b>Item Subtotal</b>
				0.00 USD
				<b>Tax</b>
				0.00 USD
				<b>Total for This Shipment</b>
				0.00 USD

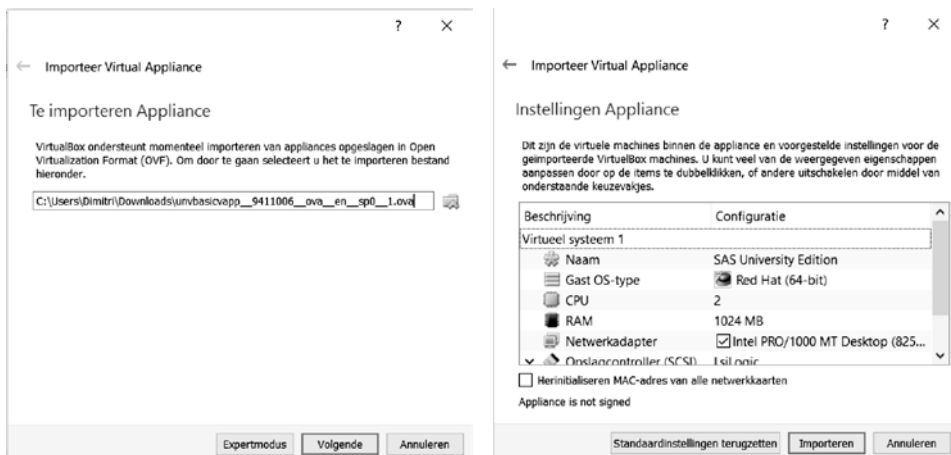
Je krijgt in je downloads map een “ova”-bestand met een lange naam:



Dat bestand bevat het volledige programma én de Linux-omgeving. De volgende stap is nu om dit programma te importeren in VirtualBox en klaar te maken om op te starten.

### 2.1.2.2 Importeren en configureren in VirtualBox

Wanneer je VirtualBox hebt geïnstalleerd, zie je geen “applicances” klaarstaan. Een appliance is een omgeving met een besturingssysteem dat de gebruiker in staat stelt virtueel te gebruiken. Je importeert de vApp van SAS UE door in het menu *Bestand* te kiezen voor *Appliance importeren*. Je gaat naar de plaats op je harde schijf waar je het ova-bestand hebt gedownload (vaak is dat in Downloads) en je selecteert het bestand zodat het kan ingevoerd worden. Het programma geeft in de volgende stap een overzicht van welke appliance je wil importeren. Om het importeren te starten, klik je op de knop *Importeren*.



Het importeren duurt even maar als VirtualBox klaar is, krijg je SAS University Edition te zien in de lijst van appliances:



Weerhoud jezelf om te dubbelklikken op het icoon! We moeten eerst nog enkele zaken configureren voordat SAS UE zal werken.

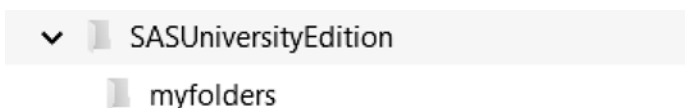
### 1. *Aanmaak mappen*

Omdat SAS UE een virtuele omgeving is, moet er een toegangspoort gemaakt worden tussen je eigen PC en de virtuele omgeving. Die toegangspoort dient dan om databestanden klaar te zetten of syntax en output te bewaren die vervolgens zichtbaar is in de virtuele omgeving.

Je maakt de volgende twee mappen aan op je computer:

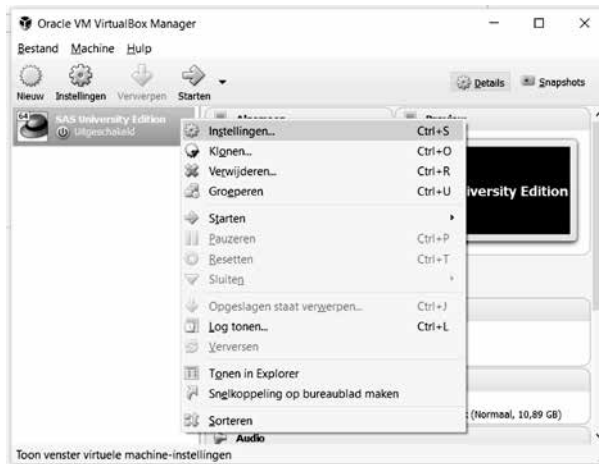
```
C:\SASUniversityEdition
C:\SASUniversityEdition\myfolders
```

Let op: je moet de naamgeving exact overnemen (Hoofdletters U en E zijn verplicht in de naam).



## 2. Instellingen van de Virtuele machine

Klik in VirtualBox rechts op de virtuele machine van SAS en kies voor *Instellingen*.

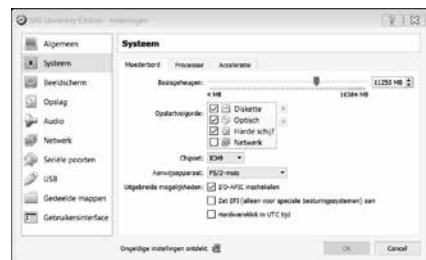


Je kan nu de standaardinstellingen wijzigen en aanpassen aan de situatie op jouw computer.

### SYSTEEM

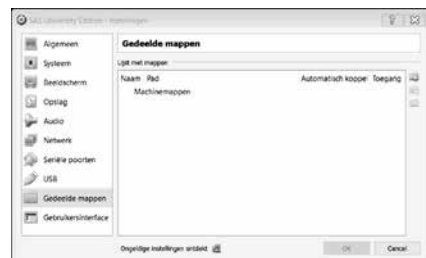
In deze optie kan je de hoeveelheid geheugen die aan SAS UE wordt toegewezen wijzigen.

Het programma zal zelf met een groene kleur aangeven hoe ver je het geheugen kan uitbreiden. Meer geheugen betekent snellere resultaten bij analyses en databeheer. De groene zone geeft aan hoe ver je kan gaan zonder de stabiliteit van je hele systeem in gevaar te brengen. Je kan tot in de rode zone gaan maar dat gaat ten koste van de stabiliteit van je computer.



### GEDEELDE MAPPEN

In deze optie koppel je de map *myfolders* aan SAS UE. Dit is noodzakelijk om SAS UE te kunnen gebruiken.



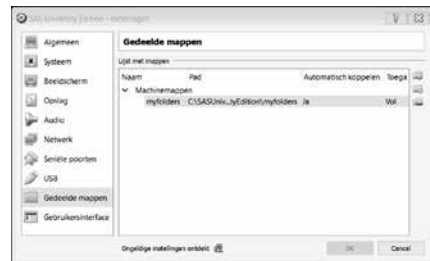
Klik op de map met het groene kruisje (📁) en voeg de map toe.

Zorg er voor dat het kruisje bij “Automatisch koppelen” ook is aangevinkt en dat het kruisje bij “Alleen-lezen” is uitgevinkt. Dat laatste is nog om in de map te kunnen schrijven. SAS moet zowel databestanden als programma’s kunnen opslaan in deze map. Als deze alleen-lezen gemaakt wordt, is dat onmogelijk en werkt het programma niet.



Als de koppeling is gelukt, dan staat myfolders nu aangegeven als een gedeelde map. Via die map zal de communicatie tussen virtuele machine en harde schijf gebeuren.

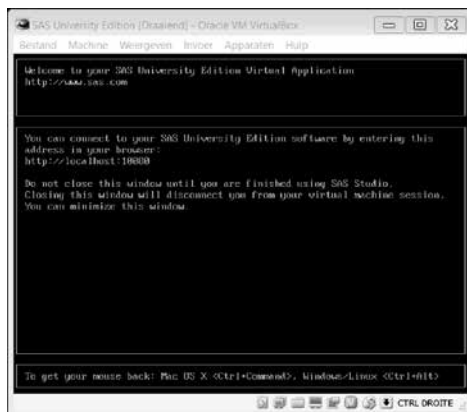
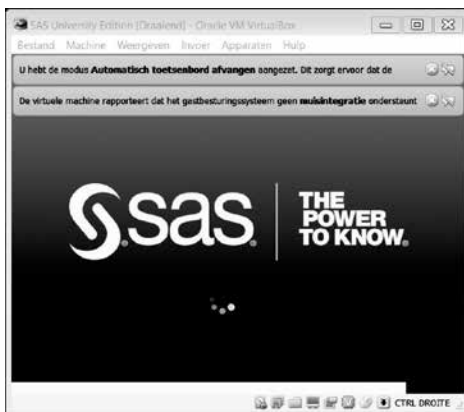
Alles wat je in deze map plaatst, zal in SAS UE zichtbaar worden.



Als deze instellingen opgeslagen zijn, ben je klaar om SAS UE op te starten.

### 2.1.2.3 SAS University Edition opstarten en gebruiken

Als alle instellingen juist staan, kan je de virtuele machine opstarten. Dat kan door te dubbelklikken op de appliance van SAS of de appliance te selecteren en op het icoon *Start* te klikken.



Na een tijdje krijg je een zwart scherm met een tekst. Dat is het teken dat SAS UE op de achtergrond draait. Je moet nu alleen het programma zichtbaar maken in een webbrowser. In het zwarte venster staat de URL waar je heen moet surfen om SAS UE te kunnen gebruiken: `http://localhost:10080`. Let op: dit kan een andere URL zijn op jouw eigen computer.

Open een webbrowser en typ dat URL-adres in de adresbalk:

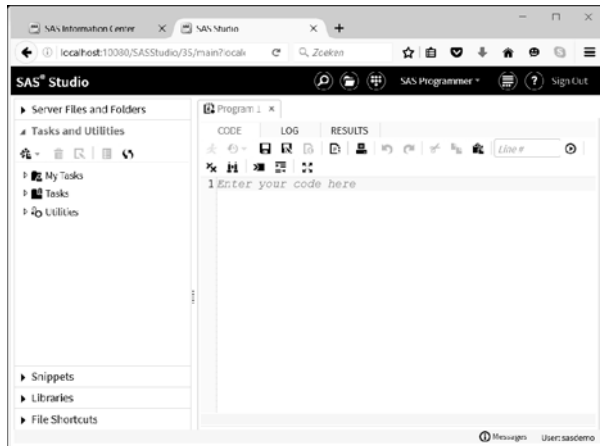


SAS UE verschijnt en je krijgt een grote knop te zien "Start SAS Studio". Nu krijg je toegang tot het programma zelf. Als in de linkerbenedenhoek "Looking for updates" verschijnt, wacht je best even tot SAS klaar is met zoeken naar updates en eventueel het installeren van de nieuwste updates, alvorens het programma op te starten.

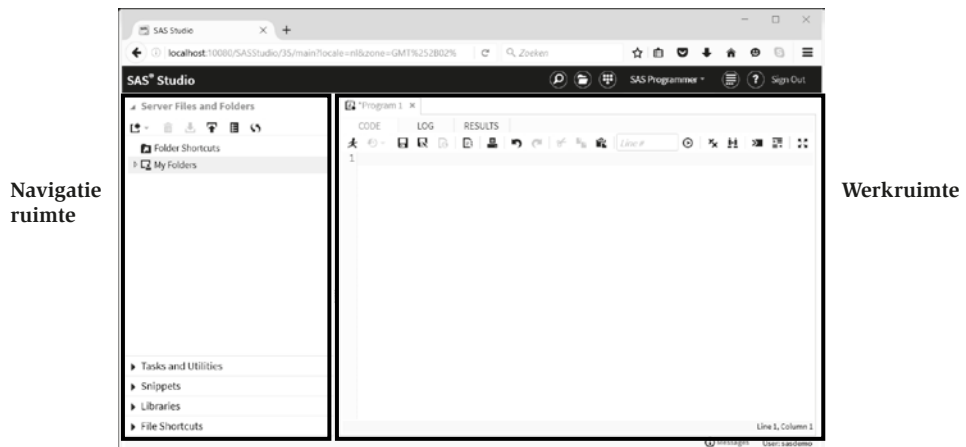
Voor je verdergaat nog een waarschuwing: het venster van VirtualBox (het zwarte scherm met de URL er in) mag je tijdens het werken met SAS UE *niet* afsluiten. Je mag het venster minimaliseren maar je mag het niet afsluiten want het venster is de toegangspoort tussen SAS en de virtuele machine. Indien je het venster afsluit terwijl SAS UE draait, kan het programma vastlopen en functioneert het nadien mogelijk niet meer.

### 2.1.3 De interface

Wanneer het programma opent, krijg je volgende interface te zien:



Het venster is ingedeeld in twee delen: het navigatiedeel aan de linkerkzijde en de werkruimte aan de rechterzijde.



### 2.1.3.1 De navigatieruimte

In de navigatieruimte staan volgende onderdelen die je bij het werken in SAS kan gebruiken:

▶ Server Files and Folders	<p>Hier heb je toegang tot de data die op je harde schijf staan. Je vindt hier de inhoud van de map <i>myfolders</i>.</p> <p>In de map <i>myfolders</i> krijg je niet alleen databestanden te zien. Naast databestanden kent SAS ook nog Catalogbestanden (extensie.sas7bcat). Catalogbestanden bevatten info die wordt gebruikt om de gegevens uit databestanden te structureren (gegevens lay-outen bv.). Hier vinden we een belangrijk verschil terug met SPSS. In SPSS worden de gegevens die nodig zijn om een gegevensbestand te structureren in één bestand met de gegevens bewaard. SAS maakt een onderscheid tussen de data zelf en het voorstellen van de data. In een databestand zit bijgevolg enkel een beschrijving van de gegevens (de <i>descriptor portion</i>) als de eigenlijke data (de <i>data portion</i>). Daarnaast is er de mogelijkheid om data naar de buitenwereld voor te stellen. Dat deel wordt opgeslagen in een apart bestand (een catalogbestand) en bevat enkel informatie over de voorstelling van gegevens. Aan de meest gebruikte catalog, de formats-catalog, wordt verder nog een hele paragraaf gewijd (zie 5.3.3).</p>
▶ Tasks and Utilities	<p>Hier krijg je voorgeprogrammeerde taken aangeboden. Je kan statistische analyses uitvoeren of grafieken maken met enkel klikken. SAS zal ook steeds de bijhorende code schrijven zodat je deze taken ook kan gebruiken om snel een bepaalde analyse in syntax te zien.</p>
▶ Snippets	<p>Snippets zijn kleine stukjes SAS-syntax die je kan gebruiken in je eigen programma's. Je kan ook eigen stukjes code bewaren als snippet en die dan later hergebruiken in je syntaxen. Snippets zijn zeer handig indien je regelmatig hetzelfde soort syntax gebruikt (bv. Voor elke tabel die je maakt is de syntax gelijkaardig en dan kan een snippet van PROC FREQ je helpen om heel snel tabellen te maken).</p>
▶ Libraries	<p>Een library is een map binnen SAS waar je data kan benaderen (zie verder). Hier krijg je een overzicht van al je libraries en de bestanden die zich daarin bevinden.</p> <p>Je vindt hier verschillende libraries die SAS steeds zal aanmaken. Sashelp bijvoorbeeld is een library waar SAS databestanden opslaat die het gebruikt bij HELP-voorbeelden. Sasuser is de standaardlibrary waar databestanden permanent worden opgeslagen en Work is de standaardlibrary waar tijdelijke databestanden opgeslagen worden. De andere libraries worden enkel in heel specifieke gevallen gebruikt. Om te kijken welke databestanden er in een library beschikbaar zijn, dubbelklik je op het icoon van de library zodat deze geopend wordt.</p>
▶ File Shortcuts	<p>Hier kan je snelle koppelingen maken naar (data)bestanden die je vaak nodig hebt.</p>






In de loop van het boek zullen we nog verder ingaan op het gebruik van de map *myfolders* en de tasks.

### 2.1.3.2 De werkruimte

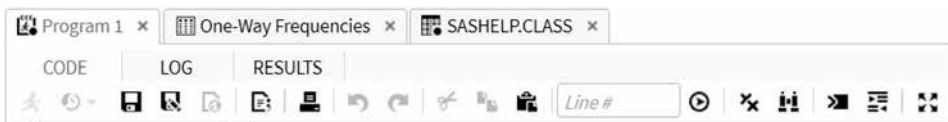
De werkruimte is georganiseerd in tabbladen. Standaard staat er één tabblad open “Program 1”.



In het lege venster eronder kan je starten met het uitwerken van SAS syntax. Je bewaart de syntax met het -icoon. Code uitvoeren doe je met het -icoon. Als je een stuk syntax selecteert, voer je ook enkel dat stuk syntax uit. Als je een stuk van je syntax wil bewaren als Snippet (stukje voorbeeldcode) dan selecteer je die code en klik je op het -icoon.

Wanneer je syntax hebt uitgevoerd, komen de resultaten in het tabblad *Results* en krijg je feedback over het uitvoeren in het tabblad *Log*. Over de concrete inhoud van dat Log-venster, zie 4.4.2.

Je kan tegelijk aan meerdere programma’s werken (al proberen we stukken syntax zo veel mogelijk samen te zetten in één programma) of je kan tasks openen en bestanden bekijken. Zodra je iets nieuws start, wordt er een extra tabblad geopend. Je kan elk van de tabbladen apart opnieuw afsluiten (klik op het kruisje aan de rechterzijde van het tabblad).

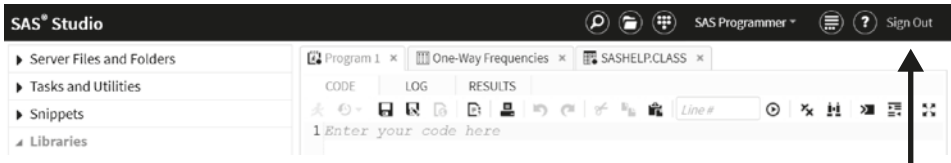


### 2.1.4 SAS UE afsluiten

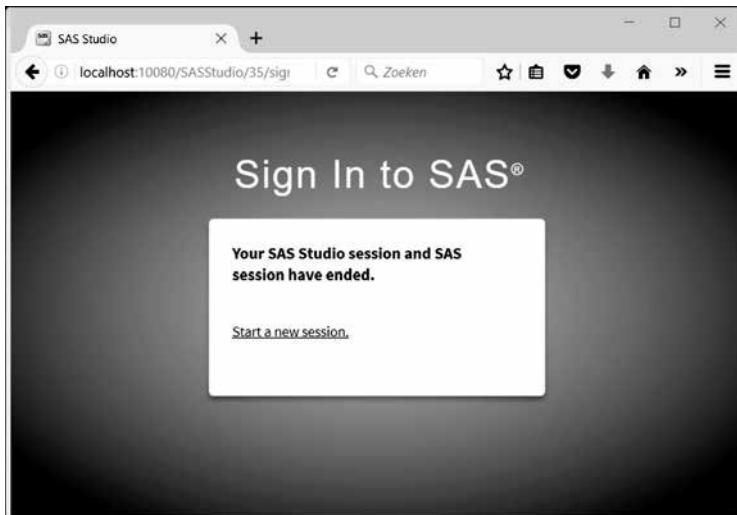
SAS UE draait als virtuele appliance op je computer. Dat betekent dat je op de juiste wijze het programma moet afsluiten om de volgende keer zonder problemen SAS UE terug te kunnen openen.

### STAP 1 SAS UE afsluiten

Eerst moet je het programma correct afsluiten *in* je browser. Dat doe je door aan de rechterzijde in de menubalk *Sign Out* te kiezen.



Als je succesvol bent uitgelogd, krijg je volgend scherm te zien:



### STAP 2 De vApp afsluiten

Vervolgens moet je de virtuele applicatie zelf ook nog afzetten. Dat gebeurt in VirtualBox. Klik rechts op de SAS appliance en kies voor *Sluiten /Uitzetten via ACPI*. ACPI staat voor *Advanced Configuration and Power Interface*. Dat is een standaard protocol om met besturingssystemen te communiceren. Wat er gebeurt als je deze optie kiest, is dat je VirtualBox een signaal laat geven naar de Linux machine die op de achtergrond draait (het zwarte scherm dat niet gesloten mocht worden). Het signaal vertelt aan Linux om zich correct af te sluiten. Op deze manier wordt de vApp gesloten en op correcte wijze uitgeschakeld.

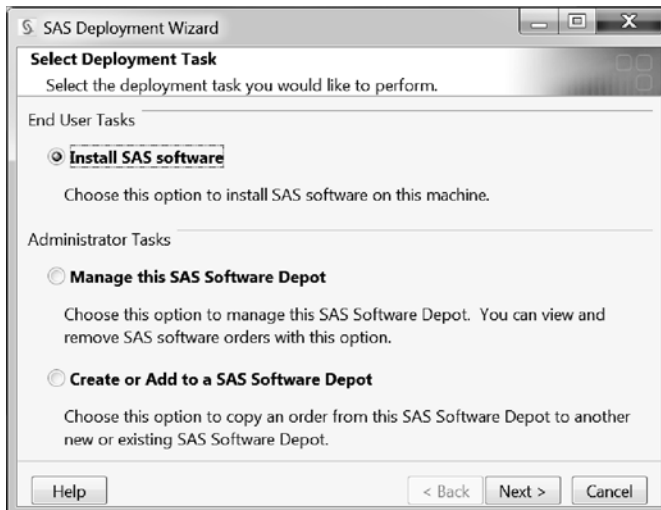
Om SAS UE opnieuw te openen, dubbelklik je de volgende keer opnieuw op de vApp van SAS in VirtualBox. Je wacht tot je een URL krijgt in het zwarte scherm en je gaat in een browser naar die URL om SAS UE opnieuw op te starten in je browser.

## 2.2 SAS Desktop

De desktop versie van SAS is een alleenstaand programma dat enkel in een Windows-versie beschikbaar is. De interface van de desktop-versie verschilt sterk<sup>1</sup> van die van SAS UE, al is het werken met SAS identiek: SAS heeft één syntaxtaal die gebruikt wordt voor datamanipulatie en analyses.

### 2.2.1 Het programma installeren

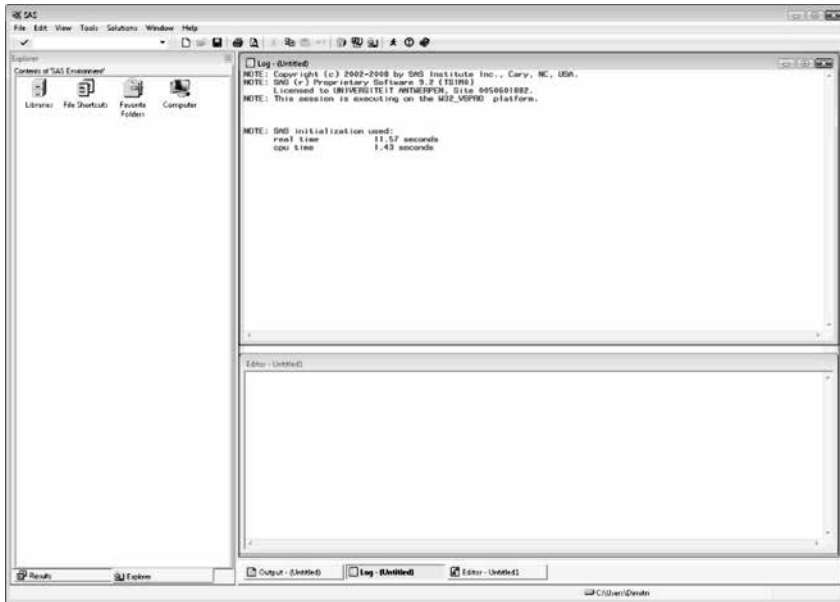
Het programma wordt geïnstalleerd op een klassieke Windows manier: je klikt op het setup.exe bestand en de installatie-wizard leid je door het hele proces.



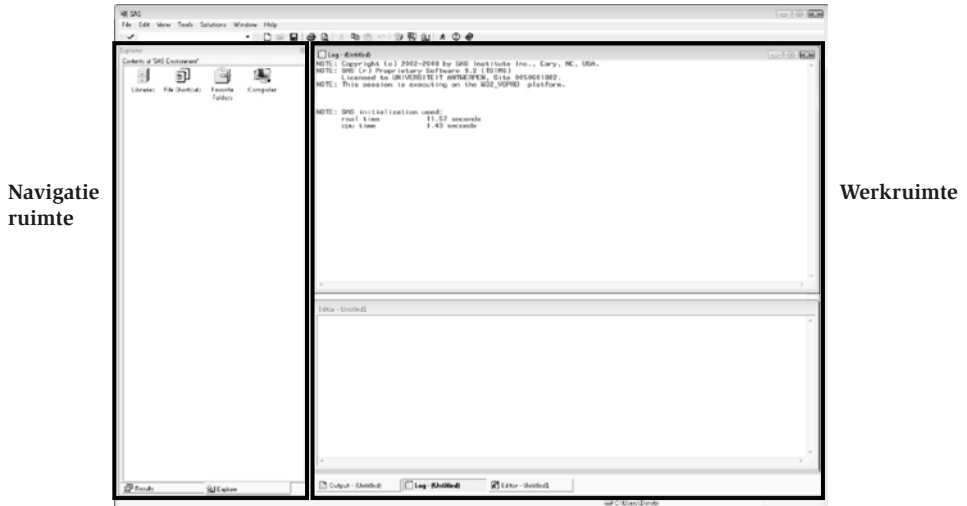
### 2.2.2 De interface

Het programma opstarten gebeurt zoals alle Windowstoepassingen: dubbelklikken op het SAS-icoon op het bureaublad of bij de programma's, SAS opzoeken en daar kiezen voor "SAS 9.4 (English)". Na een welkomtscherm, verschijnt het programma:

1. Vanaf versie 9.4 heeft de desktop versie ook een extra programma "SAS Studio". Deze SAS Studio is naar interface identiek aan de SAS UE en wordt ook via een browser uitgevoerd. Enkel VirtualBox is niet nodig voor SAS Studio. Het programma start automatisch wanneer de gebruiker het icoon dubbelklikt.



Ook al ziet de interface van de desktop versie er anders uit, het basisondercheid tussen navigatieruimte en werkruimte is ook hier aanwezig.



### 2.2.2.1 De navigatieruimte

De navigatieruimte bevat onderaan twee tabbladen: Results en Explorer. In het tabblad Results krijg je een overzichtslijst van alle analyses die je hebt uitgevoerd. Hier is verder niet veel over aan te vullen: je klikt op de analyse waarvan je het resultaat wil


zien en dat resultaat wordt in de werkruimte geopend. We richten ons daarom op het tabblad *Explorer*.

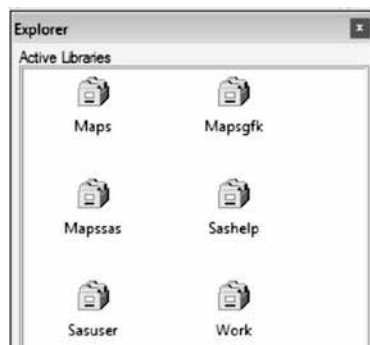
Wanneer SAS opgestart wordt, worden er in het Explorer-venster vier iconen getoond.



Dit is de betekenis van elk van deze vier iconen:

▶ Libraries	Een library is een map binnen SAS waar je data kan benaderen (zie verder). Hier krijg je een overzicht van al je libraries en de bestanden die zich daarin bevinden.
▶ File Shortcuts	Hier kan je snelle koppelingen maken naar (data)bestanden die je vaak nodig hebt.
▶ Favorite Folders	Hier kan je mappen definiëren die je regelmatig wil gebruiken.
▶ Computer	Hier krijg je toegang tot de harde schijf van je computer. Dit is vergelijkbaar met de Explorer in Windows.

We zijn in deze paragraaf voornamelijk geïnteresseerd in het bekijken van de libraries. Dubbelklik daarom op het -icoon om de libraries die SAS standaard herkent, te openen. Je merkt dat zes libraries worden aanmaakt.



Sashelp is een library waar SAS databestanden opslaat die het gebruikt bij HELP-voorbeelden. Sasuser is de standaardlibrary waar databestanden permanent worden opgeslagen en Work is de standaardlibrary waar tijdelijke databestanden opgeslagen worden. De andere libraries worden enkel in heel specifieke gevallen gebruikt (cartografie bv.). Om te kijken welke databestanden er in een library beschikbaar zijn, dubbelklik je op het icoon van de library zodat deze geopend wordt.

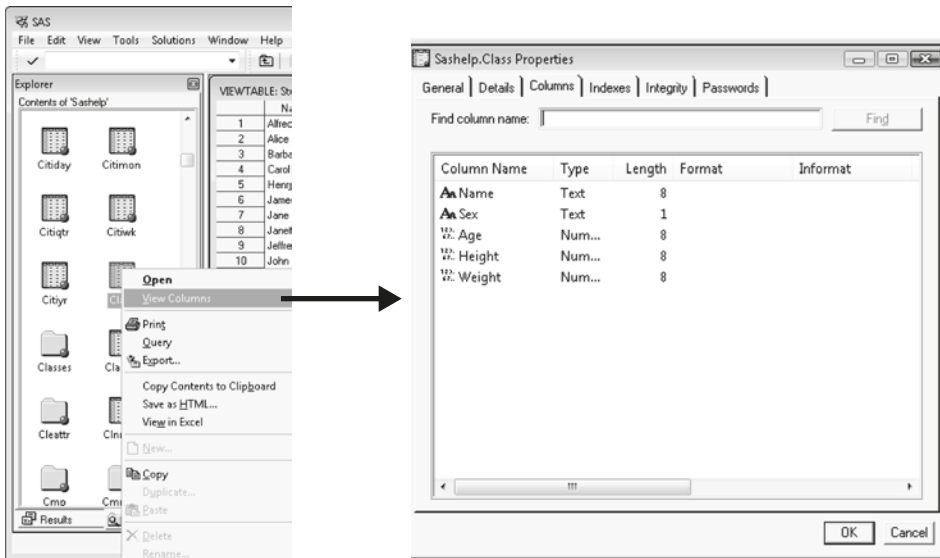
Vanuit het Explorer venster kan je vervolgens de inhoud van databestanden gaan bekijken. Dat gebeurt door te dubbelklikken op een databestand. Ga naar de library Sashelp en dubbelklik op het bestand Class. Automatisch wordt het Viewtable-venster geopend en krijg je een venster dat sterk lijkt op de Data-view in SPSS. In de kolommen staan de variabelen (in dit voorbeeld vijf), in de rijen staan de observaties (19 in totaal).

The screenshot shows the SAS software interface. On the left, the Explorer window displays the contents of the 'Sashelp' library, with the 'Class' dataset selected. The main window, titled 'VIEWTABLE: Student Data', displays a table with the following data:

	Name	Sex	Age	Height	Weight
1	Alfred	M	14	69	112.5
2	Alice	F	13	56.5	84
3	Barbara	F	13	65.3	98
4	Carol	F	14	62.8	102.5
5	Hersy	M	14	63.5	102.5
6	James	M	12	57.3	83
7	Jane	F	12	59.8	84.5
8	Janet	F	15	62.5	112.5
9	Jeffrey	M	13	62.5	84
10	John	M	12	59	99.5
11	Joyce	F	11	51.3	50.5
12	Judy	F	14	64.3	90
13	Louise	F	12	56.3	77
14	Mary	F	15	66.5	112
15	Philip	M	16	72	150
16	Robert	M	12	64.8	128
17	Ronald	M	15	67	133
18	Thomas	M	11	57.5	85
19	William	M	15	66.5	112

At the bottom of the window, a status bar indicates: 'NOTE: Table has been opened in browse mode.' The taskbar shows the user is 'C:\Users\Dimitri'.

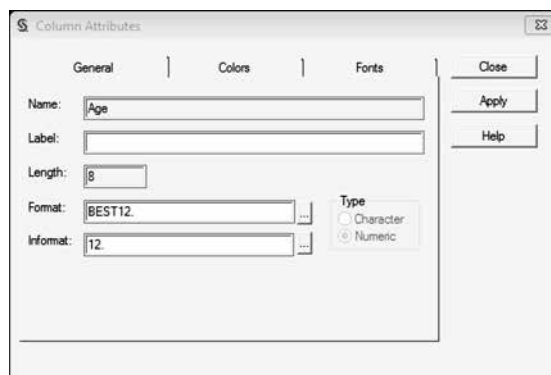
Wil je niet de data zelf zien maar de inhoud van het bestand (vergelijkbaar met de Variabele View in SPSS), dan klik je met de rechtermuisknop op een bestand en kies je voor *View Columns*. Als we dat doen bij het bestand Class, krijgen we het Properties-venster. Nu krijg je een overzicht van alle variabelen in het bestand en kan je zien welke als tekstvariabele is gedefinieerd en welke als numeriek.




Merk op dat zowel bij het bekijken van de data-inhoud als bij het bekijken van de datastructuur er aparte vensters geopend worden in de werkruimte. Je kan deze terug sluiten door te klikken op het kruisje rechtsboven in elk venster.

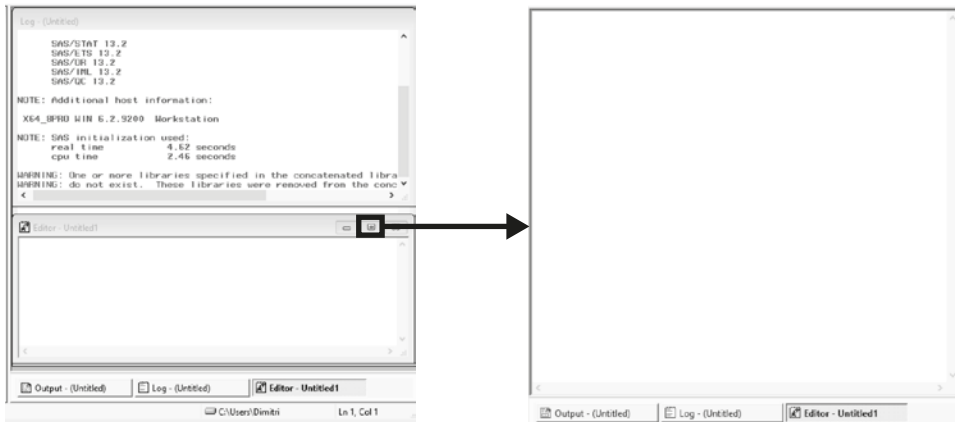


Een andere manier om de eigenschappen van een variabele te zien, is teruggaan naar het Viewtable-venster waar alle data getoond worden. Als je van een van die variabelen snel de eigenschappen wil bekijken, kan dat door rechts te klikken op die variabele en in het contextmenu te kiezen voor *Column Attributes*. SAS geeft dan een overzicht van de kenmerken van die bepaalde variabele.



### 2.2.2.2 De werkruimte

De werkruimte in SAS Desktop bevat standaard drie vensters: OUTPUT, LOG en Editor. Bij het opstarten is de werkruimte half benomen door het LOG-venster (bovenaan) en half door de Editor (onderaan). Dat werkt niet gemakkelijk dus het is best om de vensters onmiddellijk te maximaliseren via het -icoon.



Als je het venster in de werkruimte maximaliseert, heb je meer ruimte om te werken. Navigeren tussen de vensters doe je onderaan. Daar staan de namen van de verschillende vensters die geopend zijn, vermeld.

### 2.2.3 SAS Desktop afsluiten

SAS desktop sluit je af zoals elk Windows-programma: door het kruisje in de rechterbovenhoek. SAS zal nog een extra bevestiging vragen alvorens af te sluiten.

