
IBM SPSS Statistics Tips & Tricks

Edwin van de Wiel
DASC
IBM Connect 2015



***Bij deze presentatie horen
voorbeeld databestanden
en syntaxen***

Deze kunt u [hier](#) aanvragen



DASC faciliteert organisaties bij het optimaliseren van de inzet van SPSS



- IBM Advanced Business Partner
 - SPSS
 - Statistics, Modeler, Data Collection
 - Big data
- Trainingspartner
 - Avnet
 - Global Knowledge



- **Software**
 - Initieel
 - Koop, jaarhuur en maandhuur
 - Renewal Subscription & Supporting
- **Praktijktrainingen**
 - IBM SPSS trainingen
 - SPSS workshops
 - Seminars



- **Support**

- IBM Support
 - IBM Support Provider
- Gebruiker support

- **Services**

- Consultancy / meewerken
- Installatie / optimalisatie
- Applicaties





MRDM helps healthcare providers to improve quality and efficiency of medical treatments

Executive overview

Medical Research Data Management, a medical research organization in analytics and reporting solution that accelerates analysis by 99.6 percent to help insights they need to improve patient care and outcomes when it deploys IBM help from IBM Advanced Business Partner **DASC B.V.**

Groningen Regional Police Force

Gaining insight into citizen and employee satisfaction

The need:

The Groningen Regional Police Force needed to improve citizen satisfaction levels to meet the Dutch government's new targets. However, its existing paper-based satisfaction surveys were expensive to run, could not be tailored to different groups of respondents, and often failed to provide a representative sample of the local population.

The solution:

The Force worked with DASC, an IBM Business Partner, to build a sophisticated online satisfaction survey using IBM SPSS Data Collection, which automatically tailors questions to the respondent's situation – whether they are an ordinary



- **Software**
- **Praktijktrainingen**
- **Support**
- **Services**



TIPS & TRICKS



- Op verzoek van IBM
- Combinatie van uw input en onze praktijk
- Generieke en grootste gemene deler
- Demo in SPSS Statistics 23

- Na afloop ruime gelegenheid voor vragen
- Slides en materialen worden toegestuurd

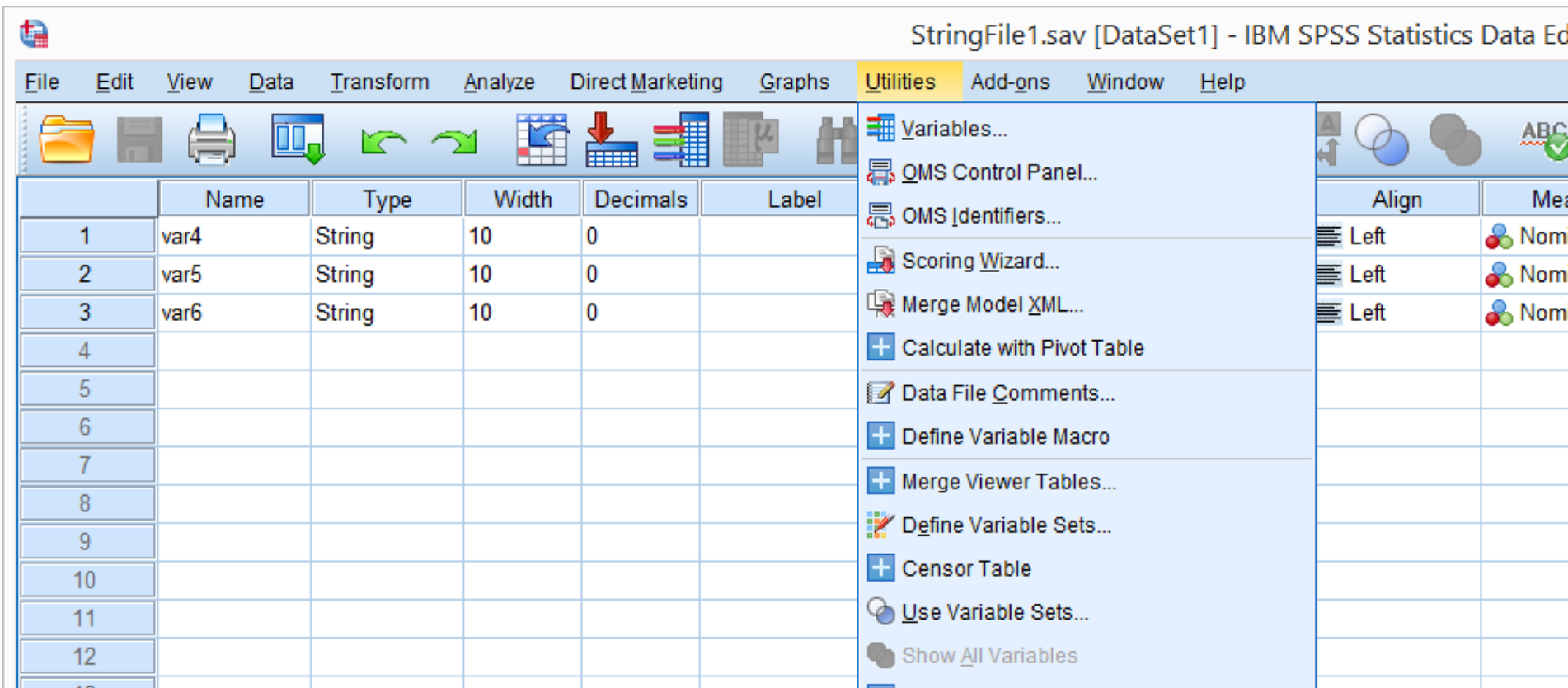


EXTENSIONS



Wat zijn Extensions

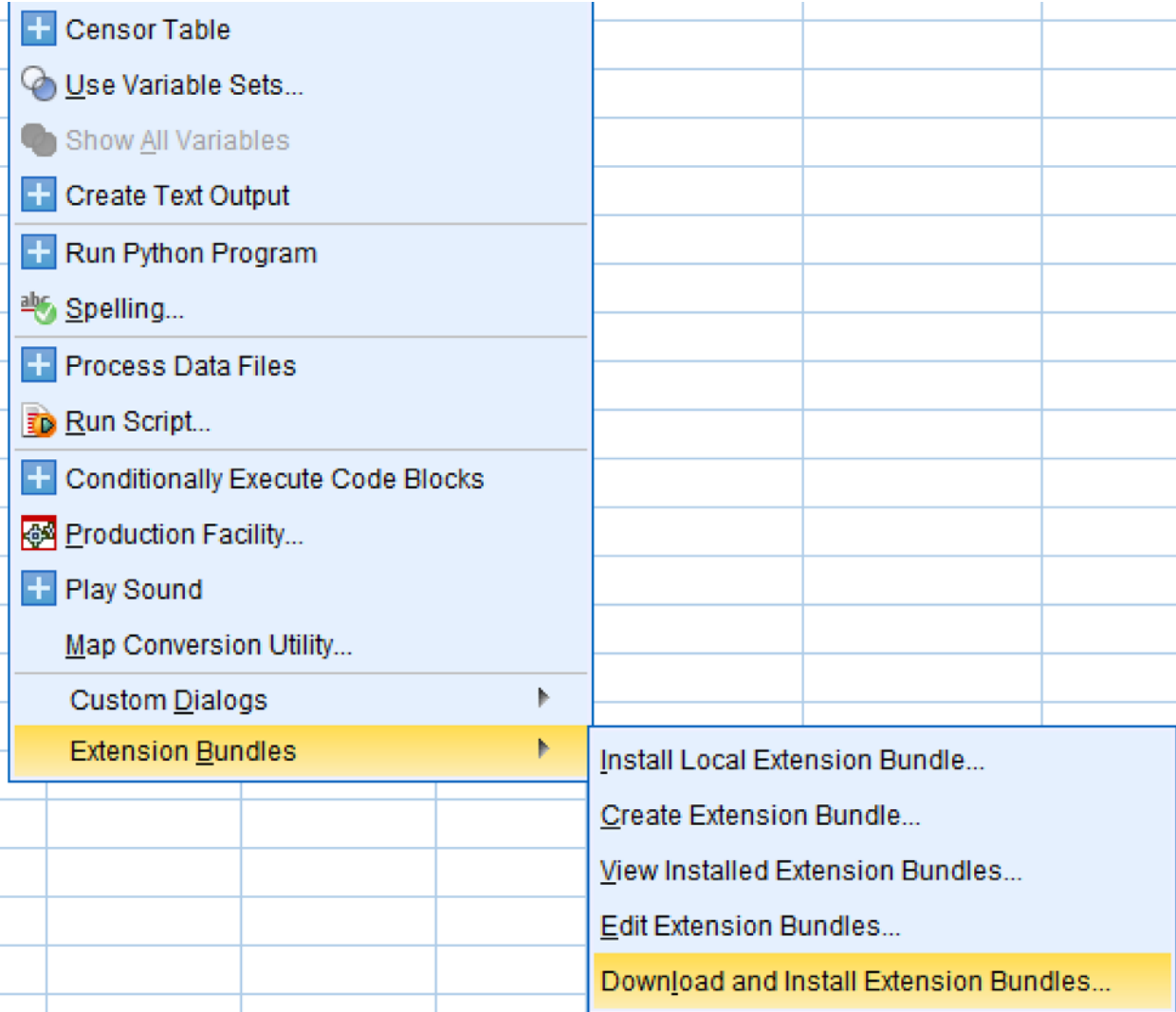
- Aanvullende functionaliteit voor SPSS
 - Werken 'hetzelfde' als standaard commando's
- Utilities → Extension Bundles



The screenshot shows the IBM SPSS Statistics Data Editor interface. The title bar indicates the file is 'StringFile1.sav [DataSet1]'. The menu bar includes File, Edit, View, Data, Transform, Analyze, Direct Marketing, Graphs, Utilities (highlighted), Add-ons, Window, and Help. The toolbar contains various icons for file operations and data manipulation. The main window displays a data table with the following columns: Name, Type, Width, Decimals, and Label. The table contains three rows of data for variables var4, var5, and var6, all of which are String type with a width of 10 and 0 decimals. The Utilities menu is open, showing options such as Variables..., QMS Control Panel..., OMS Identifiers..., Scoring Wizard..., Merge Model XML..., Calculate with Pivot Table, Data File Comments..., Define Variable Macro, Merge Viewer Tables..., Define Variable Sets..., Censor Table, Use Variable Sets..., and Show All Variables.

	Name	Type	Width	Decimals	Label
1	var4	String	10	0	
2	var5	String	10	0	
3	var6	String	10	0	
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					

Extensions



The image shows a software menu with the following items:

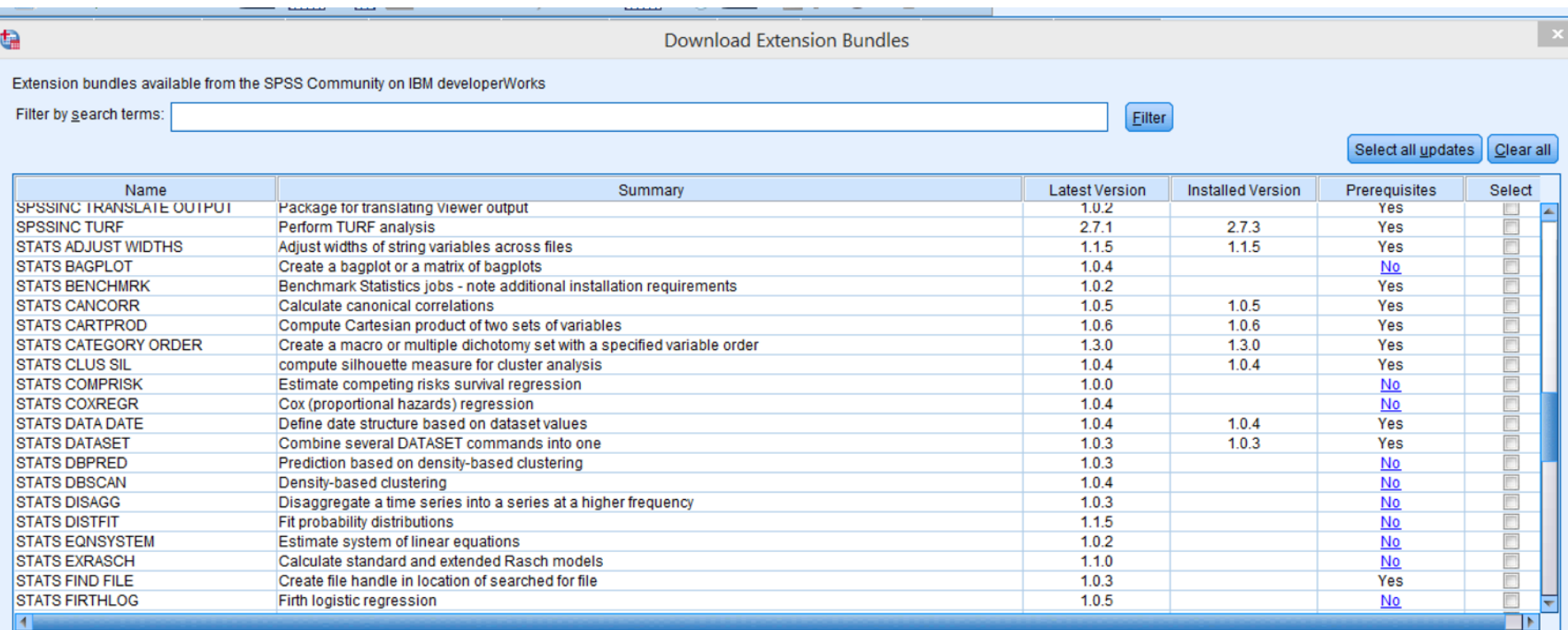
- + Censor Table
- Use Variable Sets...
- Show All Variables
- + Create Text Output
- + Run Python Program
- Spelling...
- + Process Data Files
- Run Script...
- + Conditionally Execute Code Blocks
- Production Facility...
- + Play Sound
- Map Conversion Utility...
- Custom Dialogs
- Extension Bundles

The 'Extension Bundles' sub-menu is open, showing the following options:

- Install Local Extension Bundle...
- Create Extension Bundle...
- View Installed Extension Bundles...
- Edit Extension Bundles...
- Download and Install Extension Bundles...

Extensions

- Website IBM SPSS Community
- Vanaf SPSS22 ook vanuit SPSS Statistics
- SPSS' App store



Name	Summary	Latest Version	Installed Version	Prerequisites	Select
SPSSINC TRANSLATE OUTPUT	Package for translating Viewer output	1.0.2		Yes	<input type="checkbox"/>
SPSSINC TURF	Perform TURF analysis	2.7.1	2.7.3	Yes	<input type="checkbox"/>
STATS ADJUST WIDTHS	Adjust widths of string variables across files	1.1.5	1.1.5	Yes	<input type="checkbox"/>
STATS BAGPLOT	Create a bagplot or a matrix of bagplots	1.0.4		No	<input type="checkbox"/>
STATS BENCHMRK	Benchmark Statistics jobs - note additional installation requirements	1.0.2		Yes	<input type="checkbox"/>
STATS CANCORR	Calculate canonical correlations	1.0.5	1.0.5	Yes	<input type="checkbox"/>
STATS CARTPROD	Compute Cartesian product of two sets of variables	1.0.6	1.0.6	Yes	<input type="checkbox"/>
STATS CATEGORY ORDER	Create a macro or multiple dichotomy set with a specified variable order	1.3.0	1.3.0	Yes	<input type="checkbox"/>
STATS CLUS SIL	compute silhouette measure for cluster analysis	1.0.4	1.0.4	Yes	<input type="checkbox"/>
STATS COMPRISK	Estimate competing risks survival regression	1.0.0		No	<input type="checkbox"/>
STATS COXREGR	Cox (proportional hazards) regression	1.0.4		No	<input type="checkbox"/>
STATS DATA DATE	Define date structure based on dataset values	1.0.4	1.0.4	Yes	<input type="checkbox"/>
STATS DATASET	Combine several DATASET commands into one	1.0.3	1.0.3	Yes	<input type="checkbox"/>
STATS DBPRED	Prediction based on density-based clustering	1.0.3		No	<input type="checkbox"/>
STATS DBSCAN	Density-based clustering	1.0.4		No	<input type="checkbox"/>
STATS DISAGG	Disaggregate a time series into a series at a higher frequency	1.0.3		No	<input type="checkbox"/>
STATS DISTFIT	Fit probability distributions	1.1.5		No	<input type="checkbox"/>
STATS EQNSYSTEM	Estimate system of linear equations	1.0.2		No	<input type="checkbox"/>
STATS EXTRASCH	Calculate standard and extended Rasch models	1.1.0		No	<input type="checkbox"/>
STATS FIND FILE	Create file handle in location of searched for file	1.0.3		Yes	<input type="checkbox"/>
STATS FIRTHLOG	Firth logistic regression	1.0.5		No	<input type="checkbox"/>

*Non-IBM developed extension

Extensions

Name	Summary
SETSMACRO	Paste variable set definitions into a file
Simulate Active Data	Generate simulated dataset like the active dataset
SPSSINC ANON	Anonymize variables and data
SPSSINC APRIORI	Discover frequent itemsets, association rules using the Apriori algorithm
SPSSINC BREUSCH PAGAN	Estimate a linear model and perform a heteroscedasticity test.
SPSSINC CENSOR TABLES	Censor cells of a pivot table based on the values of a test statistic
SPSSINC COMPARE DATASETS	Compare two datasets.
SPSSINC CREATE DUMMIES	Create a set of dummy variables representing the values of one or more va
SPSSINC GETURI DATA	Open an SPSS, Excel, SAS, or Stata dataset from a web url.
SPSSINC HETCOR	Calculate correlations between nominal, ordinal, and scale variables.
SPSSINC MERGE TABLES	Merge the contents of one pivot table in the Viewer into another.
SPSSINC MFP GLM	Estimate generalized fractional polynomial regression
SPSSINC MODIFY OUTPUT	Modify output titles and other aspects of objects in the Viewer
SPSSINC MODIFY TABLES	Modify the appearance of pivot tables.
SPSSINC PROCESS FILES	Apply a file of syntax to a set of data files
SPSSINC PROCESS FILES_SEARCH	Search the cases in a set of SPSS SAV files
SPSSINC PROGRAM	Extension command to run arbitrary Python programs without turning them
SPSSINC QOQLOT2	Two variable or two group Q-Q plot
SPSSINC QUANTREG	Estimate one or more conditional quantiles for a linear model.
SPSSINC RAKE	Calculate weights to control totals in up to ten dimensions by rim weighting
SPSSINC RANFOR	Estimate random forest

*Non-IBM developed extension

- Meerdere bestanden
 - Bv maandelijkse data-aanlevering in Excel
- Strings dezelfde naam maar verschillende lengtes
- Niet direct te koppelen
- Extension
 - Adjust String width across files



String variabelen samenvoegen

StringFile1.sav [DataSet1] - IBM

	Name	Type	Width	Decimals	Label	Values	Missing	Column
1	var4	String	10	0		None	None	12
2	var5	String	10	0		None	None	12
3	var6	String	10	0		None	None	12
4								
5								

StringFile2.sav [DataSet2] - IBM

	Name	Type	Width	Decimals	Label	Values	Missing	Column
1	var4	String	2	0		None	None	12
2	var5	String	6	0		None	None	12
3	var6	String	3	0		None	None	12
4								
5								
6								
7								

- In map
 - DASC\DemoGebruikersmiddag\1_Modificatie
- Databestanden
 - DASC - StringFile1.sav
 - DASC - StringFile2.sav
- Syntaxbestand
 - 1_1_Modificatie.sps
- NB: werkt alleen als de extension 'Adjust String width across files' is geïnstalleerd.

SPSS MAPS



- Sinds SPSS 20 terug van weggeweest
 - Nu in SPSS Statistics Base
- Benodigdheden
 - Kaart
 - Geografische data



- In map
 - DASC\DemoGebruikersmiddag\2_Maps
- Databestanden
 - DASC - Voorbeeld Maps.sav
- Syntaxbestand
 - nvt
- Kaart
 - DASC\DemoGebruikersmiddag\Common\Basiskaarten



SPSS STATISTICS EN PYTHON



- Universele programmeertaal
- Bijzonder geschikt voor automatiseren
- SPSS Statistics kan PYTHON script inbedden

- Krachtige manipulatie metadata
 - Metadata informatie gebruiken in script
- Doorgaans geschikter voor automatiseren dan SPSS syntax
- Veel voorbeeldscript voor SPSS op het web

- Scripts runnen
- Scripts aanpassen en combineren
- Script schrijven

```
6  
7 GET FILE='DASC - Voorbeeld Python.sav'.  
8  
9 ► BEGIN PROGRAM PYTHON.  
10  
11 print 2 + 2  
12  
13 ◀ END PROGRAM.  
14
```

- Prefix toevoegen aan alle variabelen
- Stuur e-mail bij error
- Reformat string maximaal aantal karakters
- Reformat alle date/time variabelen
- Gebruik alleen vars met minimaal significantie level



- In map
 - DASC\DemoGebruikersmiddag\3_Python
- Databestanden
 - DASC - Voorbeeld Python.sav
- Syntaxbestand
 - 3_1_Python_Eerste Python script.sps
 - 3_2_Python_Uitdraai op juiste meetniveau.sps
 - 3_3_Python_GeenLabels.sps



PERFORMANCE VERBETERING



- Meer databronnen, veel meer data
- Van PC naar thin client
- Performance staat onder druk

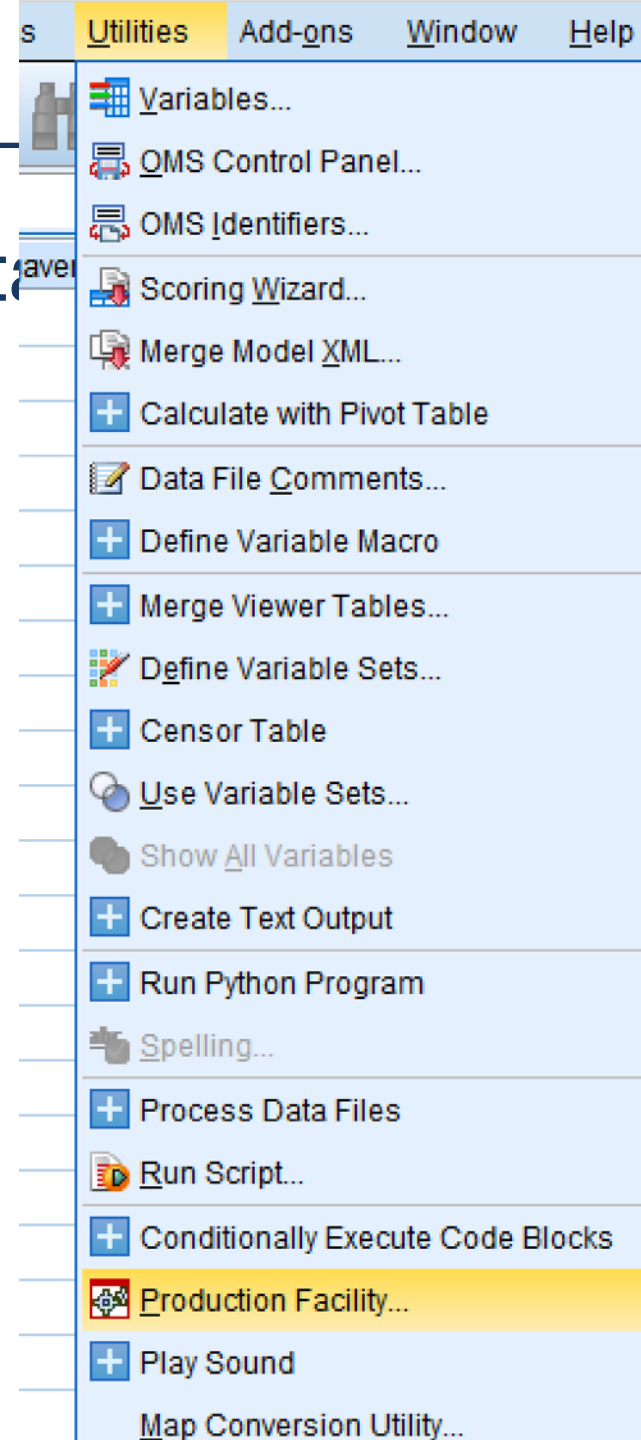
- Optimalisatie vanuit twee invalshoeken
 - Vanuit de infrastructuur & organisatie
 - Vanuit gebruiker



- Optimale PC/client
- Robuuste analytische infrastructuur
 - 64 bits
 - Van SPSS Statistics PC naar SPSS Statistics server
 - Berekeningen op andere plaats en/of andere tijd
 - Format en locatie van data
 - Slimme configuratie van netwerk
 - Afzonderlijke SPSS racebaan

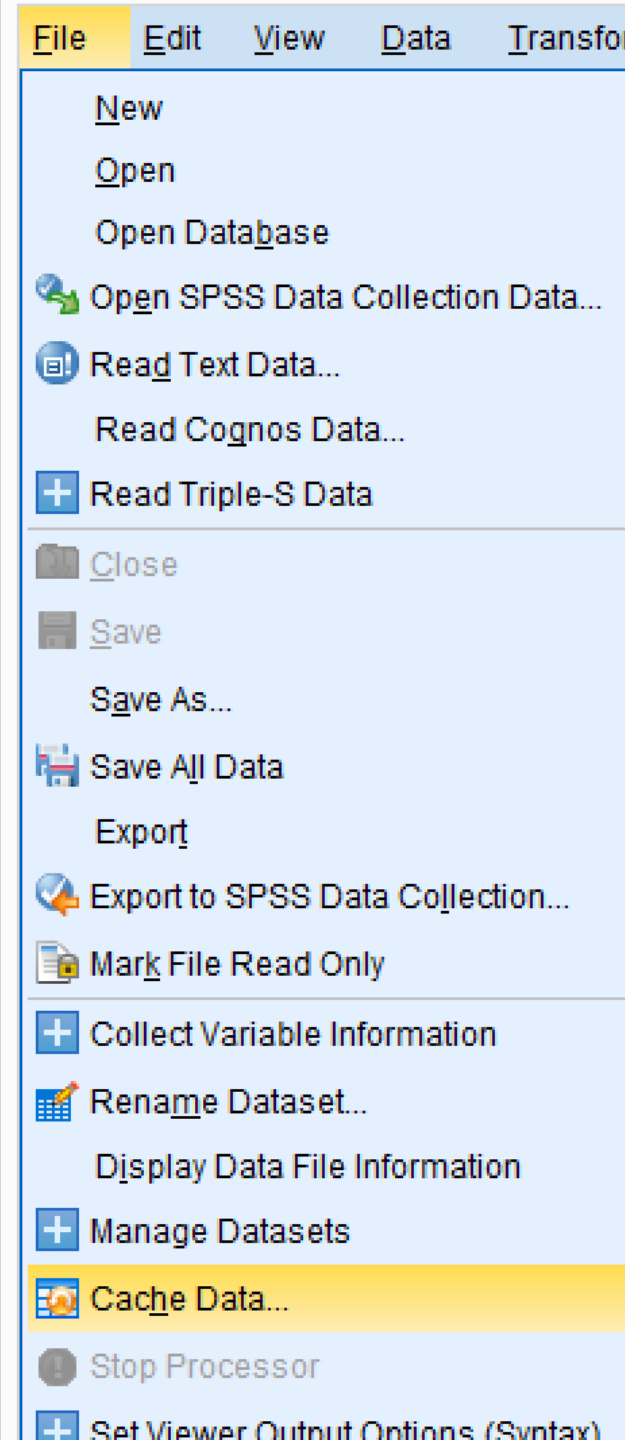
Gebruiker

- Slim omgaan met inlezen data
 - Creëren van een testdataset
 - Trekken van een steekproef
 - SAMPLE .01.
 - Reduceren aantal variabelen
 - GET FILE = 'voorbeeld.sav'
/DROP v1
- Wanneer syntax(en) runnen
 - Production facility
 - Utilities → Production Facility



Syntax optimalisatie

- Gebruik Cache
 - File → Cache Data
- Groeperen van commando's
 - Eerst alle transformaties
 - Dan analyse



- In map
 - DASC\DemoGebruikersmiddag\4_Performance
- Databestanden
 - DASC - Voorbeeld Performance.sav
- Syntaxbestand
 - 4_1_Performance_Optimalisatie_Syntax.sps



OUTPUT OP MAAT



- SPSS Statistics output
 - Jaren 70 look
 - Te veel informatie
 - Knippen tabel en plakken in Excel



- Vanaf SPSS22
 - OUTPUT MODIFY
- SPSS21 en eerder
 - Extensions
 - SPSSINC MODIFY OUTPUT



- Alleen relevante output
- In de juiste stijl
- Naar het gewenste formaat



- In map
 - DASC\DemoGebruikersmiddag\5_Output
- Databestanden
 - DASC - Voorbeeld Output.sav
- Syntaxbestand
 - 5_1_AanpassenOutput.sps

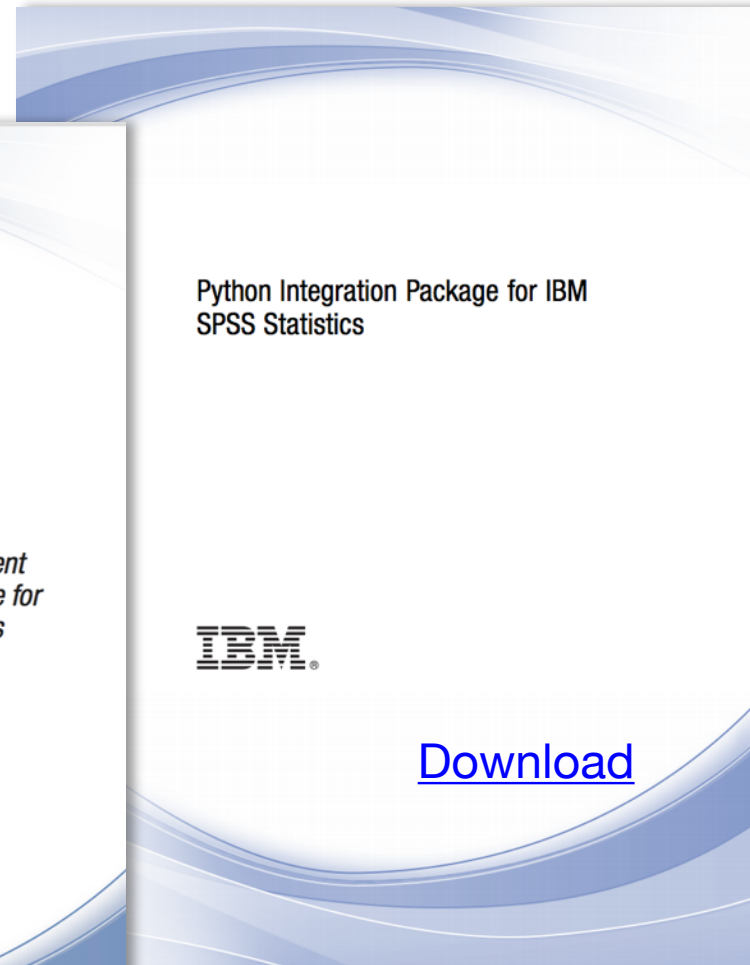
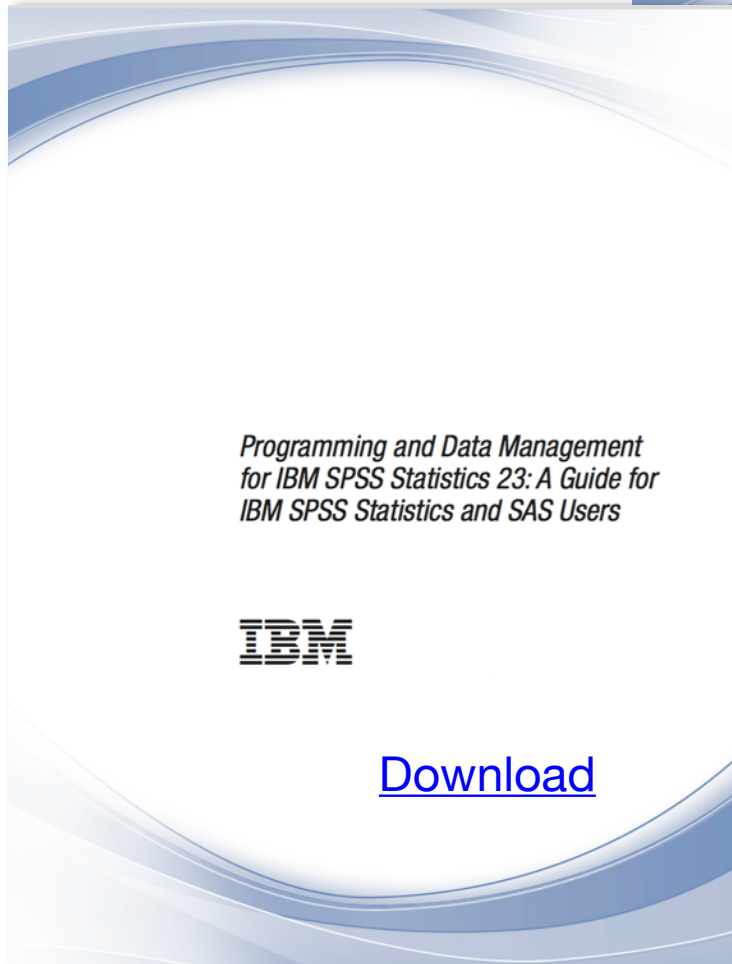


TOT SLOT



- IBM SPSS Community
 - IBM SPSS DEVELOPERWORKS
 - Onder meer:
 - SPSS Wiki
 - Blogs
 - Links





- Pragmatische en direct te gebruiken kennis
- Diverse onderwerpen en datums
 - [Kaarten maken](#) 16 juni 2015
 - [Starten met SPSS Syntax in 25 voorbeelden](#) 18 juni 2015
 - [Selecteren en hercoderen](#) 30 juni 2015
 - [Datacleaning met SPSS](#) 30 juni 2015
- Gratis workshop naar keuze bij S&S via DASC



Visschersplein 160-21
3511 LX Utrecht
085-8774406

DASC support: support@dasc.nl
DASC nieuwsbrief: [Aboneer op onze nieuwsbrief](#)
DASC events: <http://www.dasc.nl/events>

Edwin van de Wiel: evdwiel@dasc.nl
Frank Passtoors: fpasstoors@dasc.nl





www.dasc.nl



DASC: Solutions & Services for Analytical Applications

