UNIKASSEL VERSITÄT

LOEWE
Exzellente Forschung für Hessens Zukunft

translateR

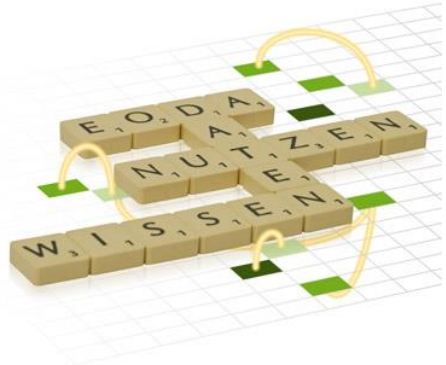
The solution for your migration processes

Oliver Bracht

July 1, 2014

About eoda

- An interdisciplinary team of data scientists, engineers, economists and social scientists,
- founded 2010 in Kassel (Germany),
- specialized in the analyzing of structured and unstructured data,
- integrated portfolio in order to solve analytical problems,
- consultation, training, customized software and services with a focus on „R“.



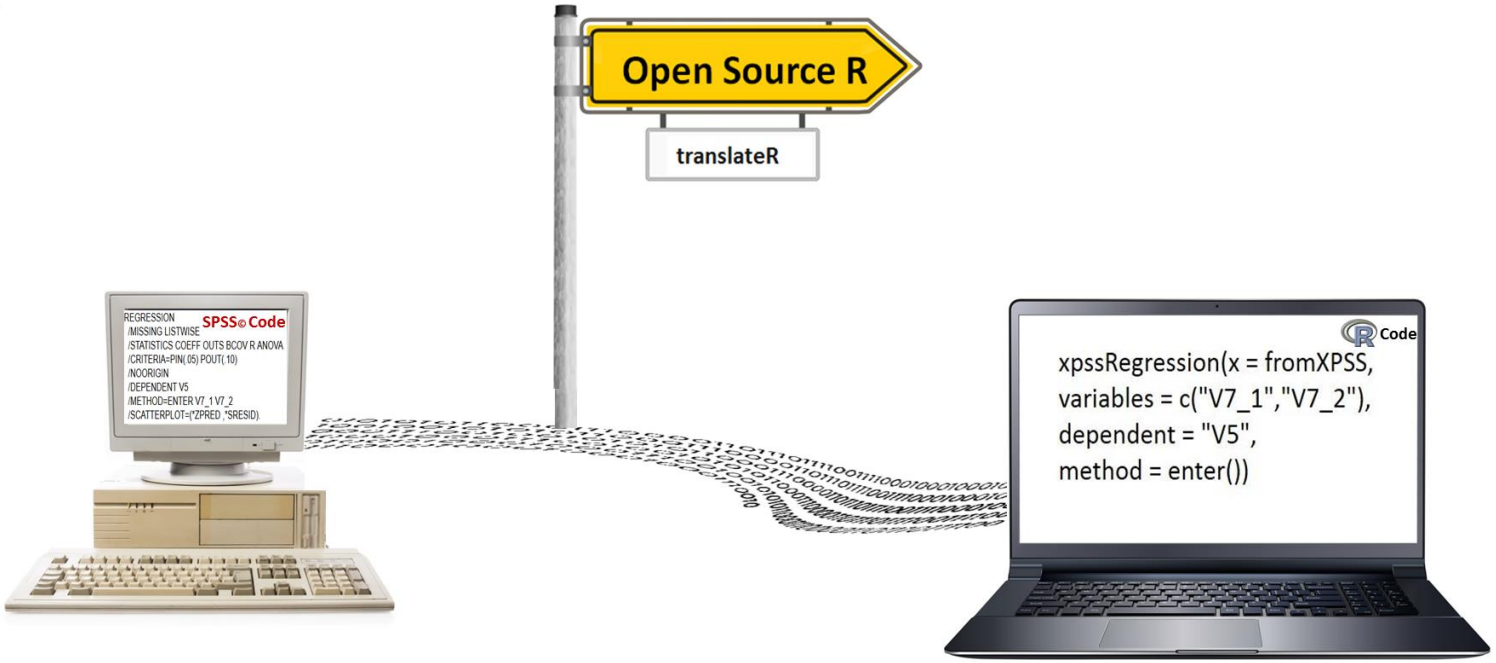
Initial situation

- SPSS®/SAS® in place
- Thousands of lines of code
- Code has grown dynamically over years
- The initial author of the code is no longer available

Migration to R

- Complexity
- Time requirements
- Costs
- Mismatch in details of concepts
- Error prown process

The solution: translateR



Fields of application

Code Migration to R

SPSS®/SAS®-like functions & features available in R

Good starting point for SPSS®/SAS® users to learn R

Quick and easy Proof Of Concept Studies

Implementation

- Cloud-based translation engine
- translateR package to run the translate code

xpssFrame-Object

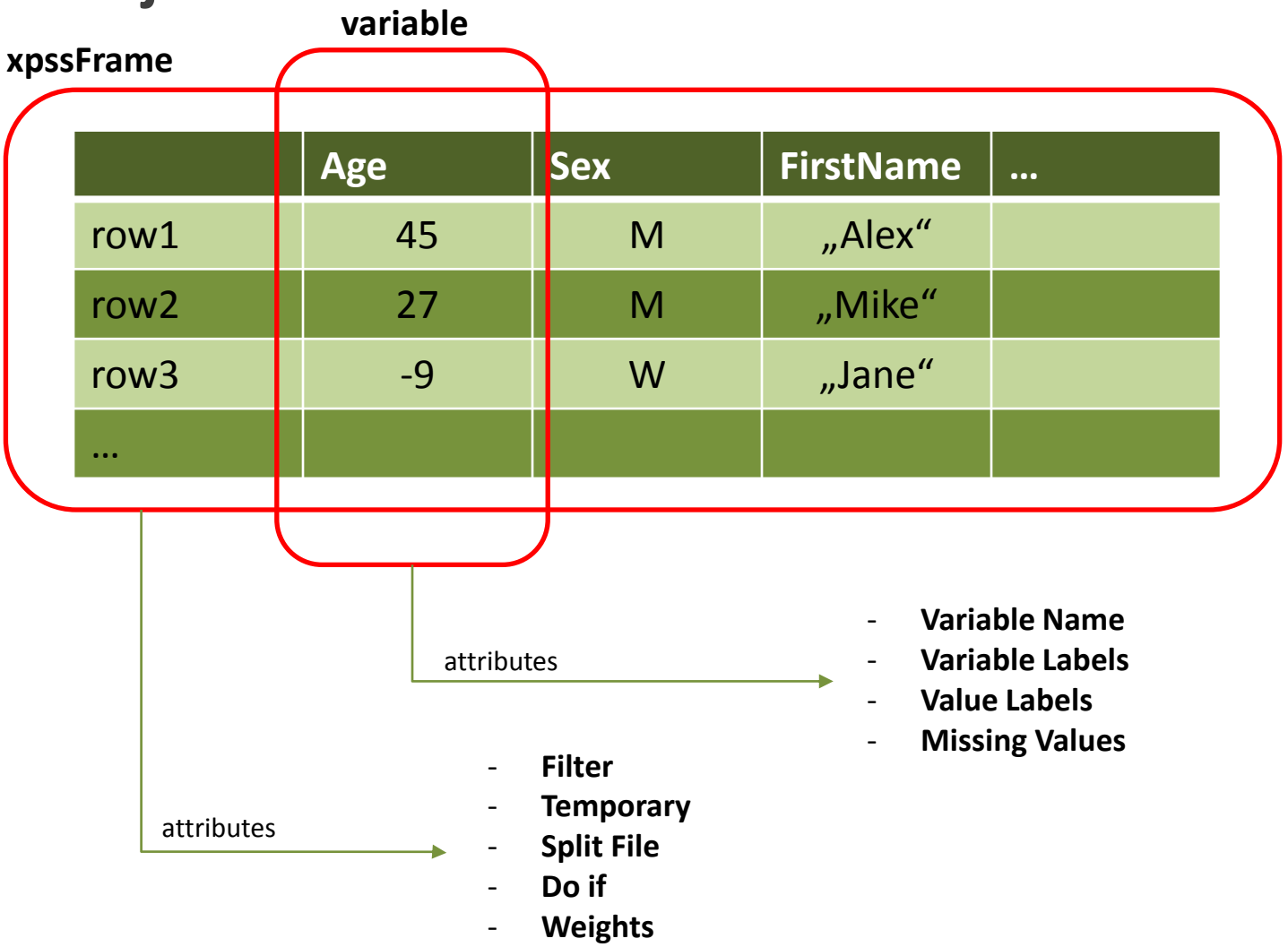
xpssFrame

	Age	Sex	FirstName	…
row1	45	M	„Alex“	
row2	27	M	„Mike“	
row3	-9	W	„Jane“	
…				

attributes

- **Filter**
- **Temporary**
- **Split File**
- **Do if**
- **Weights**

xpssFrame-Object



Naming Convention

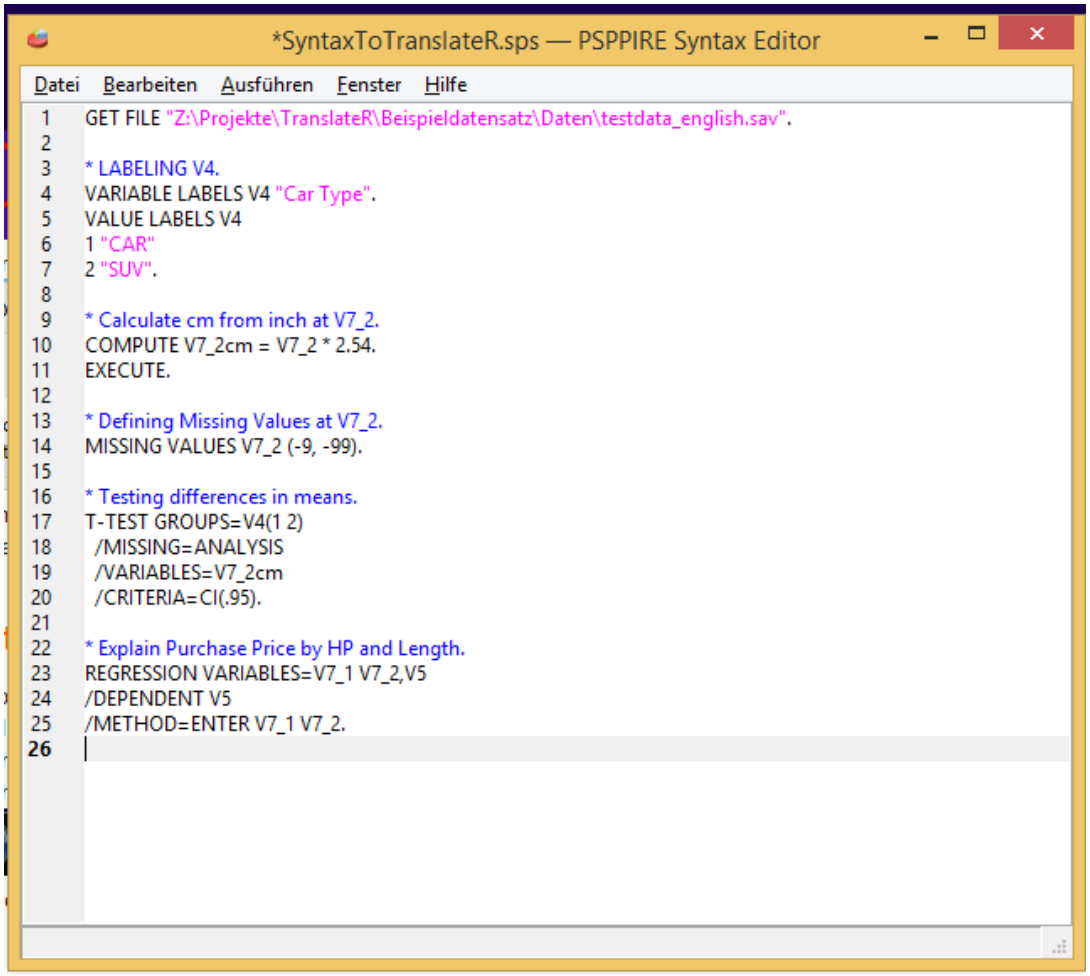
SPSS	Translation	R
VALUE LABELS	→	xpssValueLabels
MISSING VALUES	→	xpssMissingValues
…	→	xpss…

The `translateR` package provides all the funcitions to let your code run just like in SPSS

Source on Save

```
1  library(translateR)
2
3  trdat <- xpssFrame("Z:\\Projekte\\TranslateR\\Beispieldatensatz\\Daten\\testdata_english.sav")
4
5  # LABELING V4
6  trdat <- xpssVariableLabels(trdat, vars = c("V4"), labels = c("Car Type"))
7  trdat <- xpssValueLabels(trdat, vars = c("V4"), value = c(1,2), label = c("CAR","SUV"))
8
9  # Calculate cm from inch at V7_2
10 trdat$V7_2cm <- trdat$V7_2*2.54
11
12 # Defining Missing Values at V7_2
13 trdat <- xpssMissingValues(trdat, variable = c("V7_2"), as.missing = c(-9,-99))
14
15 # Testing differences in means
16 xpssT.test(trdat, t_test = "groups", groupvar = c("V4"), groups = c(1,2),
17                            vars = c("V7_2cm"), missing = "analysis", criteria = 0.95)
18
19 # Explain Purchase Price by HP and Length
20 xpssRegression(trdat, vars = c("V7_1","V7_2","V5"), dependent = c("V5"),
21                method = list(enter("V7_1","V7_2")))
```

Demo



translateR

TranslateR

TranslateR translates SPSS-Code to R-Code automatically.

This is alpha software. Things might work as expected but most of the time they won't. Send bug reports and feedback to our support team. (support@eoda.de)

SPSS

Put your SPSS code here...

Translate »

R

Output will be here...

TranslateR version: 0.5.53-SNAPSHOT, WebClient version: 0.3.-1-SNAPSHOT, SDMLib version: 1.0 - Build: 30-06-2014

translateR

TranslateR

TranslateR translates SPSS-Code to R-Code automatically.

This is alpha software. Things might work as expected but most of the time they won't. Send bug reports and feedback to our support team. (support@eoda.de)

SPSS

```
GET FILE "Z:\Projekte\TranslateR\Beispieldatensatz\Daten\testdata_english.sav".

* LABELING V4.
VARIABLE LABELS V4 "Car Type".
VALUE LABELS V4
1 "CAR"
2 "SUV".

* Calculate cm from inch at V7_2.
COMPUTE V7_2cm = V7_2 * 2.54.
EXECUTE.

* Defining Missing Values at V7_2.
MISSING VALUES V7_2 (-9, -99).

* Testing differences in means.
T-TEST GROUPS=V4(1 2)
  /MISSING=ANALYSIS
  /VARIABLES=V7_2cm
  /CRITERIA=CI(.95).

* Explain Purchase Price by HP and Length.
REGRESSION VARIABLES=V7_1 V7_2,V5
/DEPENDENT V5
/METHOD=ENTER V7_1 V7_2.
```

Translate »

R

Output will be here...

TranslateR version: 0.5.53-SNAPSHOT, WebClient version: 0.3.-1-SNAPSHOT, SDMLib version: 1.0 - Build: 30-06-2014

translateR

eoda
daten ▸ wissen ▸ nutzen

TranslateR

TranslateR translates SPSS-Code to R-Code automatically.

This is alpha software. Things might work as expected but most of the time they won't. Send bug reports and feedback to our support team. (support@eoda.de)

SPSS

```
GET FILE "Z:\Projekte\TranslateR\Beispieldatensatz\Daten\testdata_english.sav".

* LABELING V4.
VARIABLE LABELS V4 "Car Type".
VALUE LABELS V4
1 "CAR"
2 "SUV".

* Calculate cm from inch at V7_2.
COMPUTE V7_2cm = V7_2 * 2.54.
EXECUTE.

* Defining Missing Values at V7_2.
MISSING VALUES V7_2 (-9, -99).

* Testing differences in means.
T-TEST GROUPS=V4(1 2)
  /MISSING=ANALYSIS
  /VARIABLES=V7_2cm
  /CRITERIA=CI(.95).

* Explain Purchase Price by HP and Length.
REGRESSION VARIABLES=V7_1 V7_2,V5
/DEPENDENT V5
/METHOD=ENTER V7_1 V7_2.
```

Translate »

R

```
trdat <- xpssFrame("Z:\\Projekte\\TranslateR\\Beispieldatensatz\\Daten\\testdata_english.sav")

# LABELING V4

trdat <- xpssVariableLabels(trdat, vars = c("V4"), labels = c("Car Type"))

trdat <- xpssValueLabels(trdat, vars = c("V4"), value = c(1,2), label = c("CAR","SUV"))

# Calculate cm from inch at V7_2

trdat$V7_2cm <- trdat$V7_2*2.54

# Defining Missing Values at V7_2

trdat <- xpssMissingValues(trdat, variable = c("V7_2"), as.missing = c(-9,-99))

# Testing differences in means

xpssT.test(trdat, t_test = "groups", groupvar = c("V4"), groups = c(1,2), vars = c("V7_2cm"), missing = "analysis", criteria = 0.95)

# Explain Purchase Price by HP and Length

xpssRegression(trdat, vars = c("V7_1","V7_2","V5"), dependent = c("V5"), method = list(enter("V7_1","V7_2")))
```

TranslateR version: 0.5.53-SNAPSHOT, WebClient version: 0.3.-1-SNAPSHOT, SDMLib version: 1.0 - Build: 30-06-2014

eoda
daten ▸ wissen ▸ nutzen

TranslateR

Copy

TranslateR translates SPSS-Code to R-Code automatically.

This is alpha software. Things might work as expected but most of the time they won't. Send bug reports and feedback to our support team. (support@eoda.de)

SPSS

```
GET FILE "Z:\Projekte\TranslateR\Beispieldatensatz\Daten\testdata_english.sav".

* LABELING V4.
VARIABLE LABELS V4 "Car Type".
VALUE LABELS V4
1 "CAR"
2 "SUV".

* Calculate cm from inch at V7_2.
COMPUTE V7_2cm = V7_2 * 2.54.
EXECUTE.

* Defining Missing Values at V7_2.
MISSING VALUES V7_2 (-9, -99).

* Testing differences in means.
T-TEST GROUPS=V4(1 2)
  /MISSING=ANALYSIS
  /VARIABLES=V7_2cm
  /CRITERIA=CI(.95).

* Explain Purchase Price by HP and Length.
REGRESSION VARIABLES=V7_1 V7_2,V5
/DEPENDENT V5
/METHOD=ENTER V7_1 V7_2.
```

Translate

R

```
trdat <- xpssFrame("Z:\\Projekte\\TranslateR\\Beispieldatensatz\\Daten\\testdata_english.sav")

# LABELING V4

trdat <- xpssVariableLabels(trdat, vars = c("V4"), labels = c("Car Type"))

trdat <- xpssValueLabels(trdat, vars = c("V4"), value = c(1,2), label = c("CAR","SUV"))

# Calculate cm from inch at V7_2

trdat$V7_2cm <- trdat$V7_2*2.54

# Defining Missing Values at V7_2

trdat <- xpssMissingValues(trdat, variable = c("V7_2"), as.missing = c(-9,-99))

# Testing differences in means

xpssT.test(trdat, t_test = "groups", groupvar = c("V4"), groups = c(1,2), vars = c("V7_2cm"), missing = "analysis", criteria = 0.95)

# Explain Purchase Price by HP and Length

xpssRegression(trdat, vars = c("V7_1","V7_2","V5"), dependent = c("V5"), method = list(enter("V7_1","V7_2")))
```

TranslateR version: 0.5.53-SNAPSHOT, WebClient version: 0.3.-1-SNAPSHOT, SDMLib version: 1.0 - Build: 30-06-2014

UNIKASSEL VERSITÄT

LOEWE
Exzellente Forschung für Hessens Zukunft

For more information about translateR and the registration for the Beta Test visit

http://www.eoda.de/en/translateR.html

Ludwig-Erhard-Straße 8
34131 Kassel
Germany

+49 (0) 561/202724-40

http://www.eoda.de
http://blog.eoda.de
http://twitter.com/datennutzen

info@eoda.de