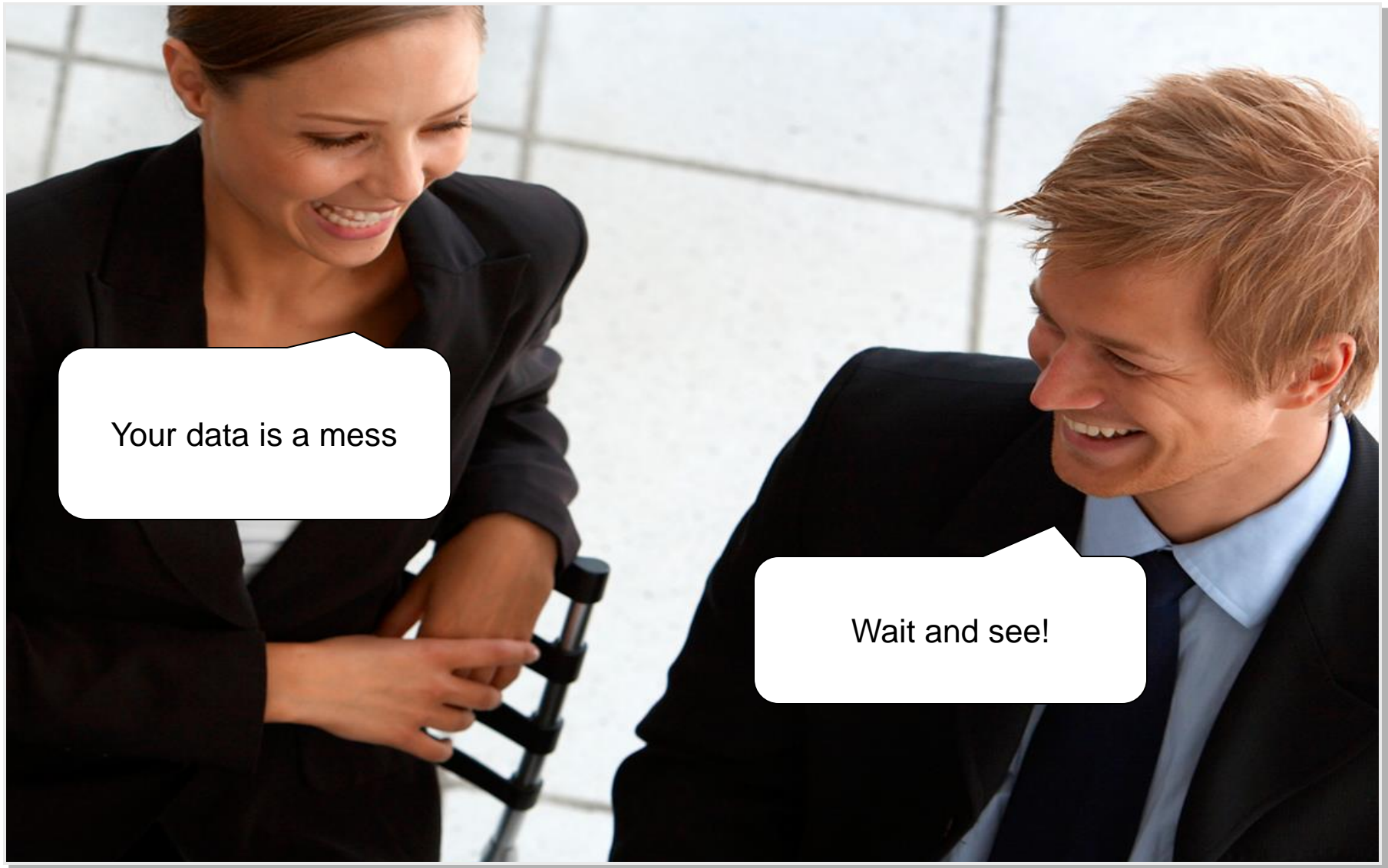


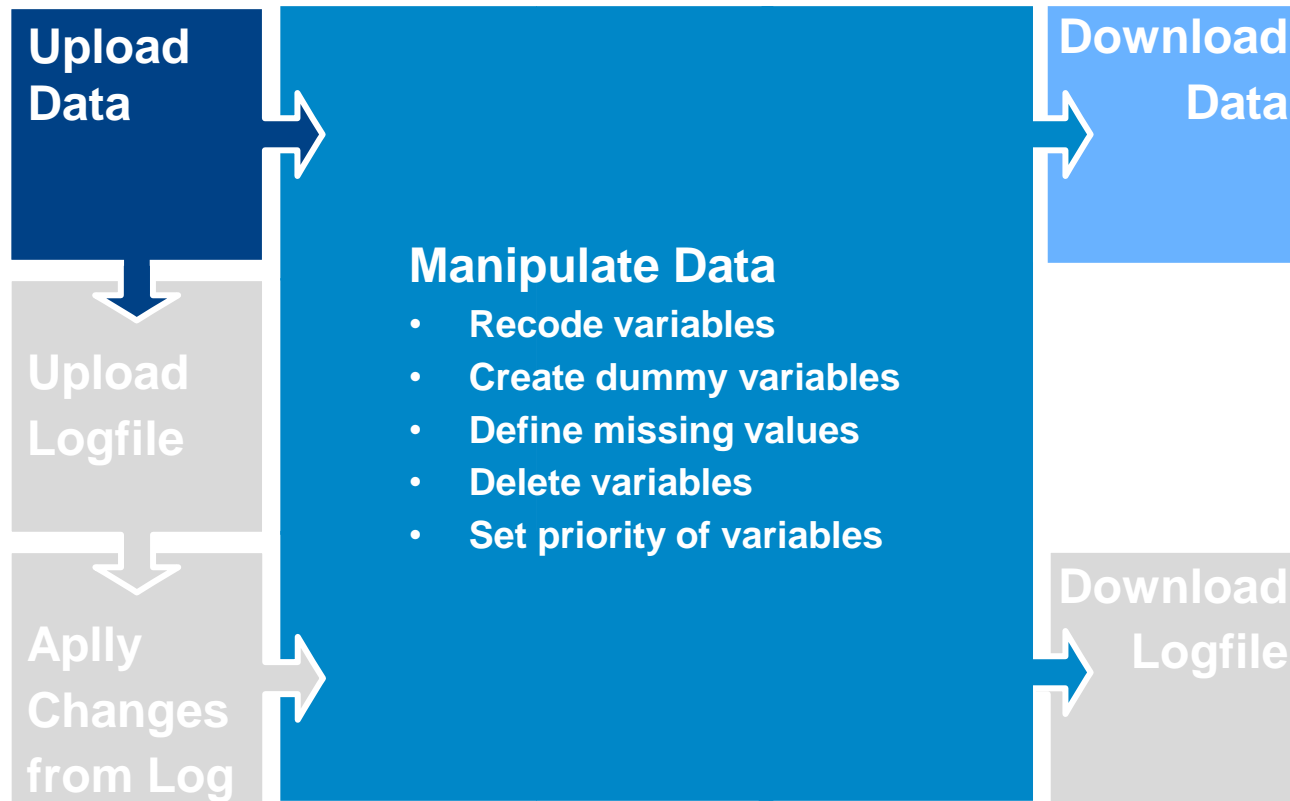
Data Input Preparation Tool – Manual



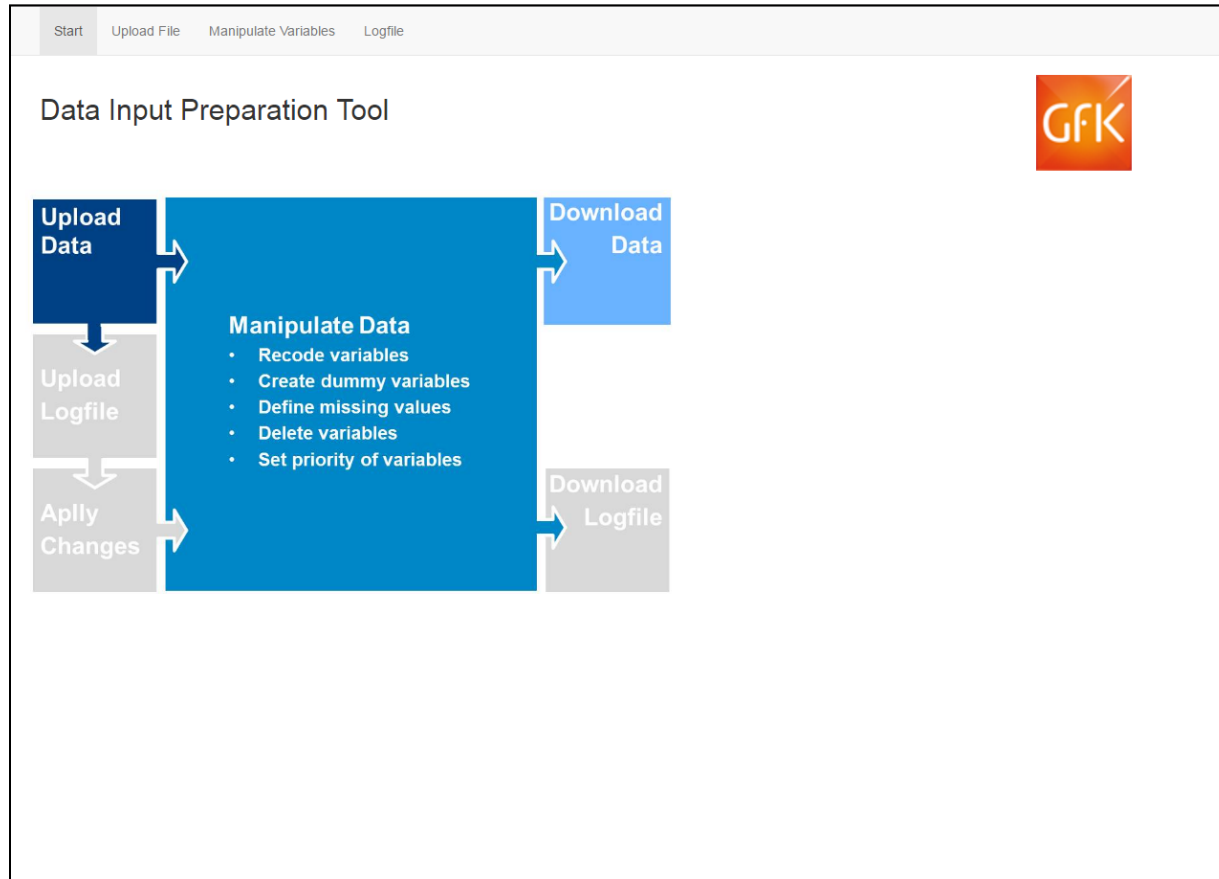
Outline

	Slide
1. Workflow	X
2. The App	X
2.1 Start	X
2.2 Upload a File	X
2.3 Manipulate Variables	X
2.3.1 Recode	X
2.3.2 Create Dummy Variables	X
2.3.3 Define Missing Values	X
2.3.4 Delete Variables	X
2.3.5 Set Priority of a Variable	X
2.4 Use a Logfile	X
3. The Logfile	X

Workflow



Data Input



Data Input

Start
Upload File
Manipulate Variables
Logfile

Select sav-File

Choose SPSS File

Browse...
No file selected

Download

Upload your sav-file here

Later, you can download the revised file here.

Data Input



Start Upload File Manipulate Variables Logfile

Select sav-File

Choose SPSS File

Browse... ExampleData.sav Upload complete.

Download

Trim data?

Delete columns: uniform cells

Delete columns: unique cells

ID variable

Please select ID variable

no ID variable

continue

	s06	s07	s08
2	2	46	3
3	3	24	1
4	4	44	1
5	5	27	2
6	6	24	1
7	7	45	3
8	8	26	3
9	9	46	2
10	10	45	1

Showing 1 to 10 of 12 entries

Previous 1 2 Next

es Logfile

ID variable

Please select ID variable

no ID variable

ID

s03

s04

s05

s06

s07

s08

46	3	-1
24	1	1
44	1	1
27	2	

When you have uploaded a sav-file, select the ID variable of your dataset. This variable will kept untouched by all manipulations. If there is no ID variable or you do not want to define one, simply choose ,no ID variable'.

Data Input

Start
Upload File
Manipulate Variables
Logfile

Select sav-File

Choose SPSS File

Browse... ExampleData.sav

Upload complete

Download

Trim data?

Delete columns: uniform cells

Delete columns: unique cells

Show 10 entries

	ID	s03	s04	s05	s06	s07	s08
1	1	45	2	2	1	1	0
2	2	46	3	-1	1	3	1
3	3	24	1	1	1	3	5
4	4	44	1	1	1	4	0
5	5	27	2		1	1	15
6	6	24	1		1	4	7
7	7	45	3	3	1	3	8
8	8	26	3	-99	1	1	3
9	9	46	2	2	1	1	0
10	10	45	1	1	1	3	12

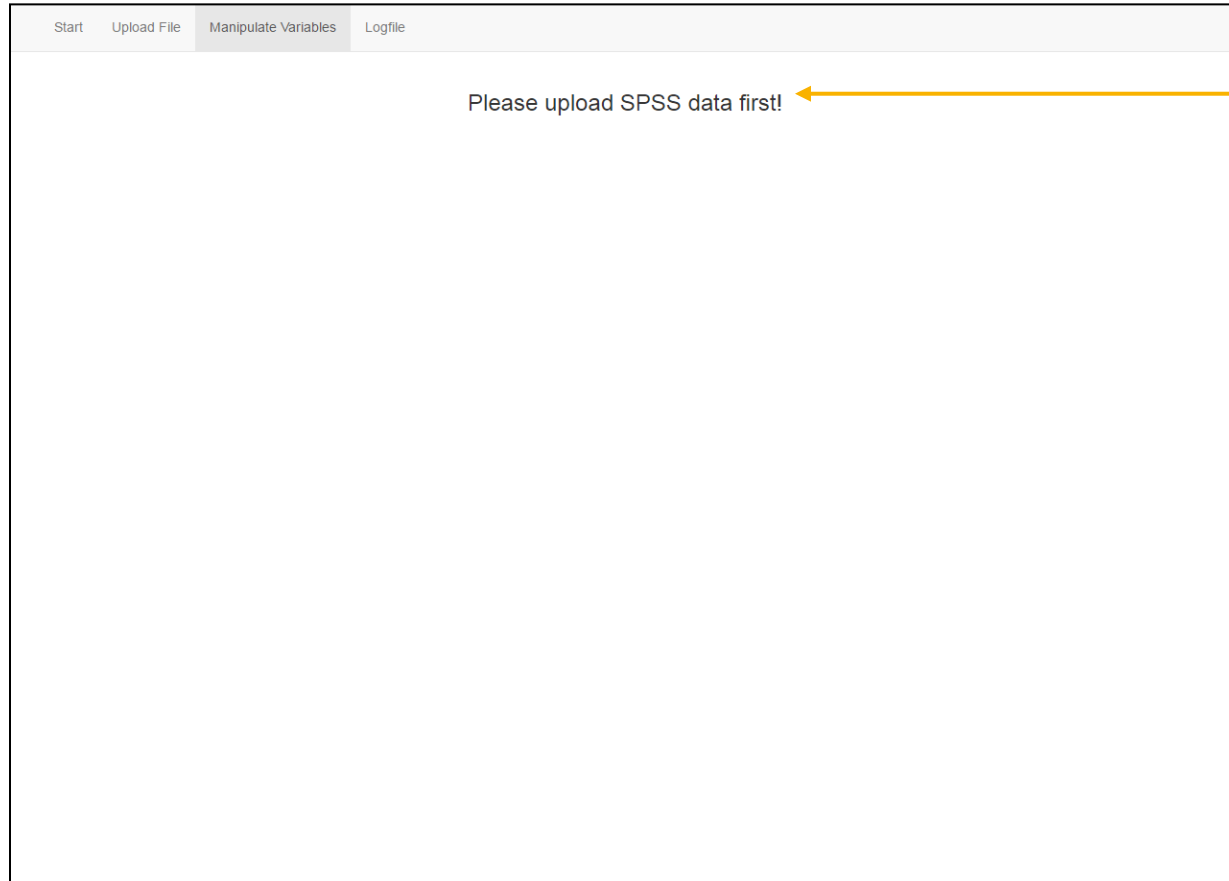
Showing 1 to 10 of 12 entries

Previous 1 2 Next

This gives you an overview of your dataset. Note that changes carried out by the app will not appear here until you upload your data again.

Here, you can delete all variables that have the same entry, like variable 's06' (→ **Delete columns: uniform cells**), or you can delete all variables that differ in all cells, like 'ID' (→ **Delete columns: unique cells**; Of course, it makes no sense to delete the ID variable. Therefore, you can make variables untouchable. See the previous slide for details.)

Data Input



Oops! Seems like you have to upload your data first.

Data Input

Start
Upload File
Manipulate Variables
Logfile

Manipulate Variables

Go to variable

Delete variable:

Yes No

Create dummy variable:

Yes No

Priority

☐ low ☐ mid ☐ high

Missing Values

Non-Responder

new

Recode!

Filter

new

Recode!

SQSD

new

Recode!

Other

new

Recode!

ID

Responder ID

recode 1 (1) to

recode 2 (2) to

recode 3 (3) to

recode 4 (4) to

recode 5 (5) to

recode 6 (6) to

recode 7 (7) to

recode 8 (8) to

recode 9 (9) to

recode 10 (10) to

recode 11 (11) to

recode 12 (12) to

The header of the current variable.

Navigate through your variables.

Delete this variable (see slide X).

Create dummy variables for this variable (see slide X).

Recode different missing values in this variable or in all variables (see slide X).

Set the priority of this variable (see slide X).

Recode single values in this variable (see slide X).

Manipulate Variables

Recode Variables

- Recode values in single variables

App

s03

In welchem Bundesland wohnen Sie?

recode 44 (Baden-Württemberg) to

1

recode 45 (Bayern) to

2

recode 46 (Berlin) to

3

recode 23 (Brandenburg) to

4

recode 24 (Bremen) to

5

recode 26 (Hamburg) to

6

recode 27 (Hessen) to

7

Recode!

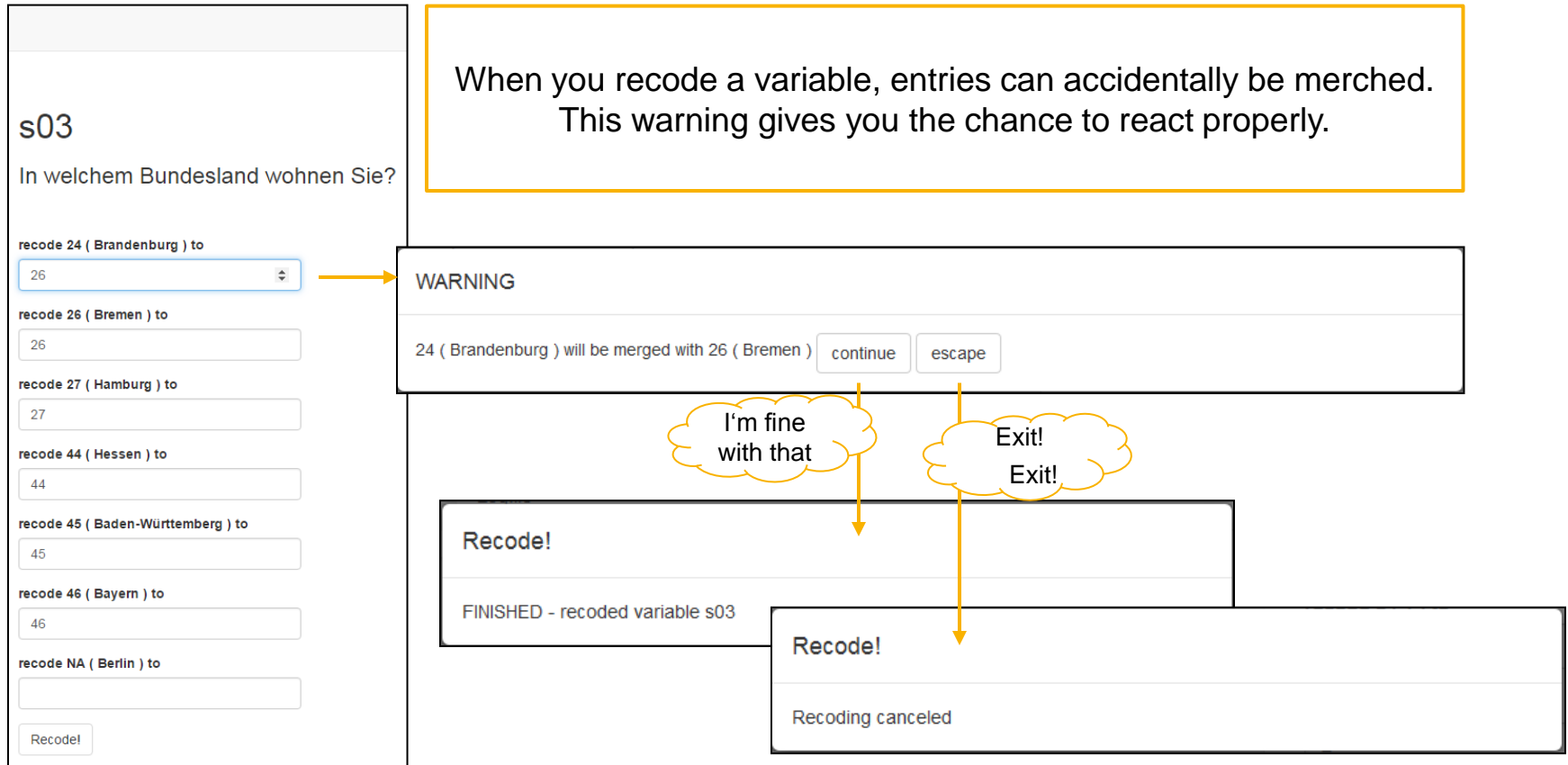
Data

s03	s03
45	2
46	3
24	5
44	1
27	7
24	5
45	2
26	6
46	3
45	2
27	7
...	...

Logfile

Variable	Level	oldvalue	newvalue
s03	Baden-Württemberg	44	1
s03	Bayern	45	2
s03	Berlin	46	3
s03	Brandenburg	23	4
s03	Bremen	24	5
s03	Hamburg	26	6
s03	Hessen	27	7
...

Data Input



Manipulate Variables

Create Dummy Variable

- Create a dummy variables for a specific variable
- Saved in Variable „Create Dummy“ in logfile (1 = yes; 0 = no)

App

Create dummy variable:

YesNo

Data

s04	s04__1	s04__2	s04__3
2	0	1	0
3	0	0	1
1	1	0	0
1	1	0	0
2	0	1	0
1	1	0	0
3	0	0	1
3	0	0	1
2	0	1	0
1	1	0	0
3	0	0	1
...	...		

Logfile

Variable	...	Create Dummy
s03	...	0
s04	...	1
s04	...	1
s04	...	1
s05	...	0
s05	...	0
s05	...	0
...

Manipulate Variables

Define Missing Values

- Recode values in the complete data set or in single variables
- Recode different missing values with different meanings

App
Data
Logfile

Missing Values

Non-Responder

new

Filter

new

SQSD

new

Other

new

↑

Recode missing values in the complete data set.

↑

Recode missing values in current variable.

s05	s05
2	2
-1	-66
1	1
1	1
-	-77
-	-77
3	3
-99	-99
2	2
1	1
3	3
...	...

➔

Variable	...	MV1 _old	MV1 _new	MV2 _old
s05	...	-1	-66	-
s05	...	na	na	na
s05	...	na	na	na
s05	...	na	na	na
s05	...	na	na	na
s05	...	na	na	na
s05	...	na	na	na
s05	...	na	na	na
...

Manipulate Variables

Delete Variables

- Delete unnecessary variables
- Saved in Variable „Remain“ in logfile (1 = yes; 0 = no)

App

Delete variable:

Yes

No

Data

s06	s07	s08
1	1	0
1	3	1
1	3	5
1	4	0
1	1	15
1	4	7
1	3	8
1	1	3
1	1	0
1	3	12
1	4	1
...		

s06	s08
1	0
1	1
1	5
1	0
1	15
1	7
1	8
1	3
1	0
1	12
1	1
...	

Logfile

Variable	...	Remain
s06	...	1
s07	...	0
s07	...	0
s07	...	0
s08	...	1
s08	...	1
s08	...	1
...

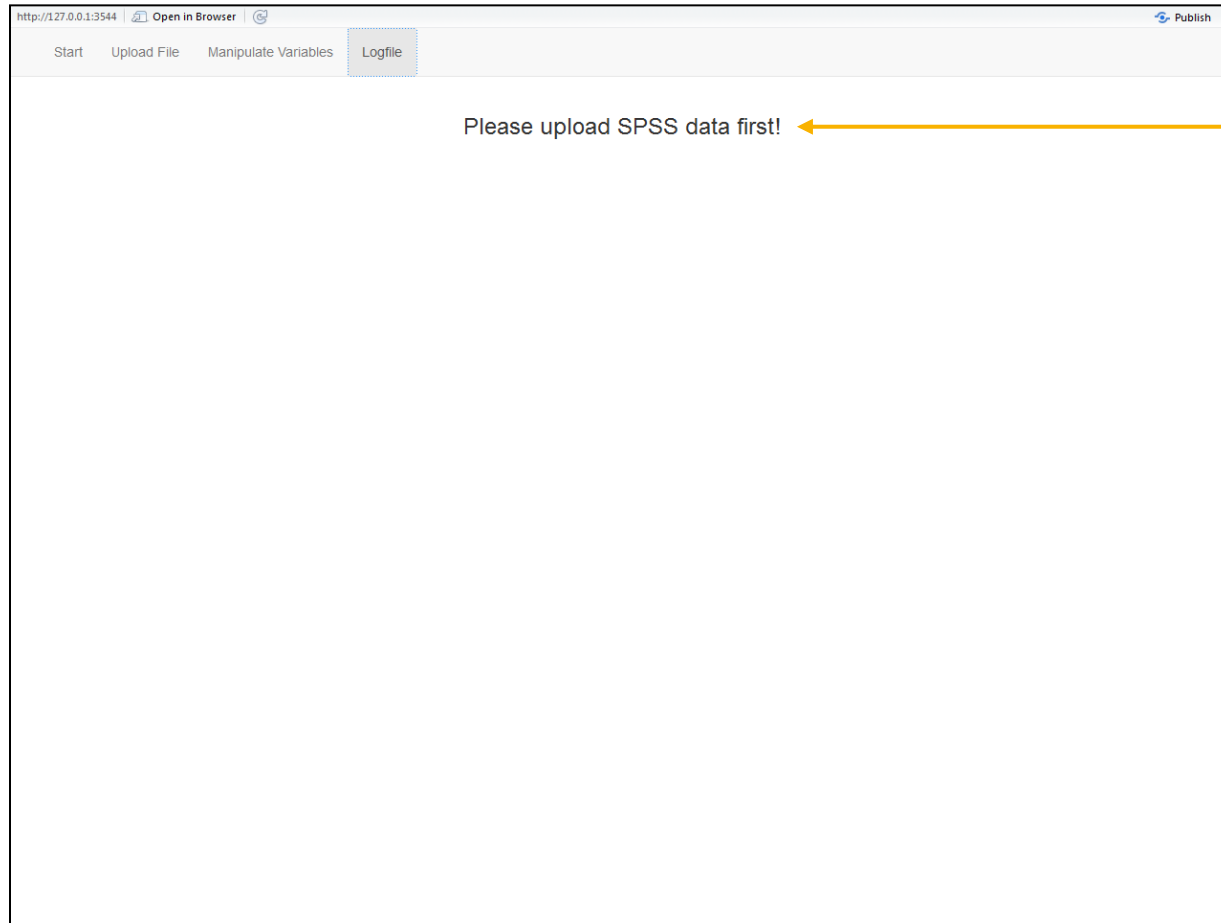
Manipulate Variables

Set Priority of Variable

- For further purposes, the priority of single variables can be defined
- Saved in Variable „Priority“ in logfile (1 = low; 2 = mid; 3 = high)
- Default is „low“.

App	Data	Logfile																											
<div> Priority <input type="radio"/> low <input type="radio"/> mid <input checked="" type="radio"/> high </div>	No changes will occur in data.	<table border="1"> <thead> <tr> <th>Variable</th> <th>...</th> <th>Priority</th> </tr> </thead> <tbody> <tr> <td>s03</td> <td>...</td> <td>1</td> </tr> <tr> <td>s04</td> <td>...</td> <td>3</td> </tr> <tr> <td>s04</td> <td>...</td> <td>3</td> </tr> <tr> <td>s04</td> <td>...</td> <td>3</td> </tr> <tr> <td>s05</td> <td>...</td> <td>1</td> </tr> <tr> <td>s05</td> <td>...</td> <td>1</td> </tr> <tr> <td>s05</td> <td>...</td> <td>1</td> </tr> <tr> <td>...</td> <td>...</td> <td>...</td> </tr> </tbody> </table>	Variable	...	Priority	s03	...	1	s04	...	3	s04	...	3	s04	...	3	s05	...	1	s05	...	1	s05	...	1
Variable	...	Priority																											
s03	...	1																											
s04	...	3																											
s04	...	3																											
s04	...	3																											
s05	...	1																											
s05	...	1																											
s05	...	1																											
...																											

Data Input



Oops! Seems like you have to upload your data first.

Data Input

http://127.0.0.1:3544 | Open in Browser | Publish

Start Upload File Manipulate Variables **Logfile**

Select Logfile

Upload Logfile

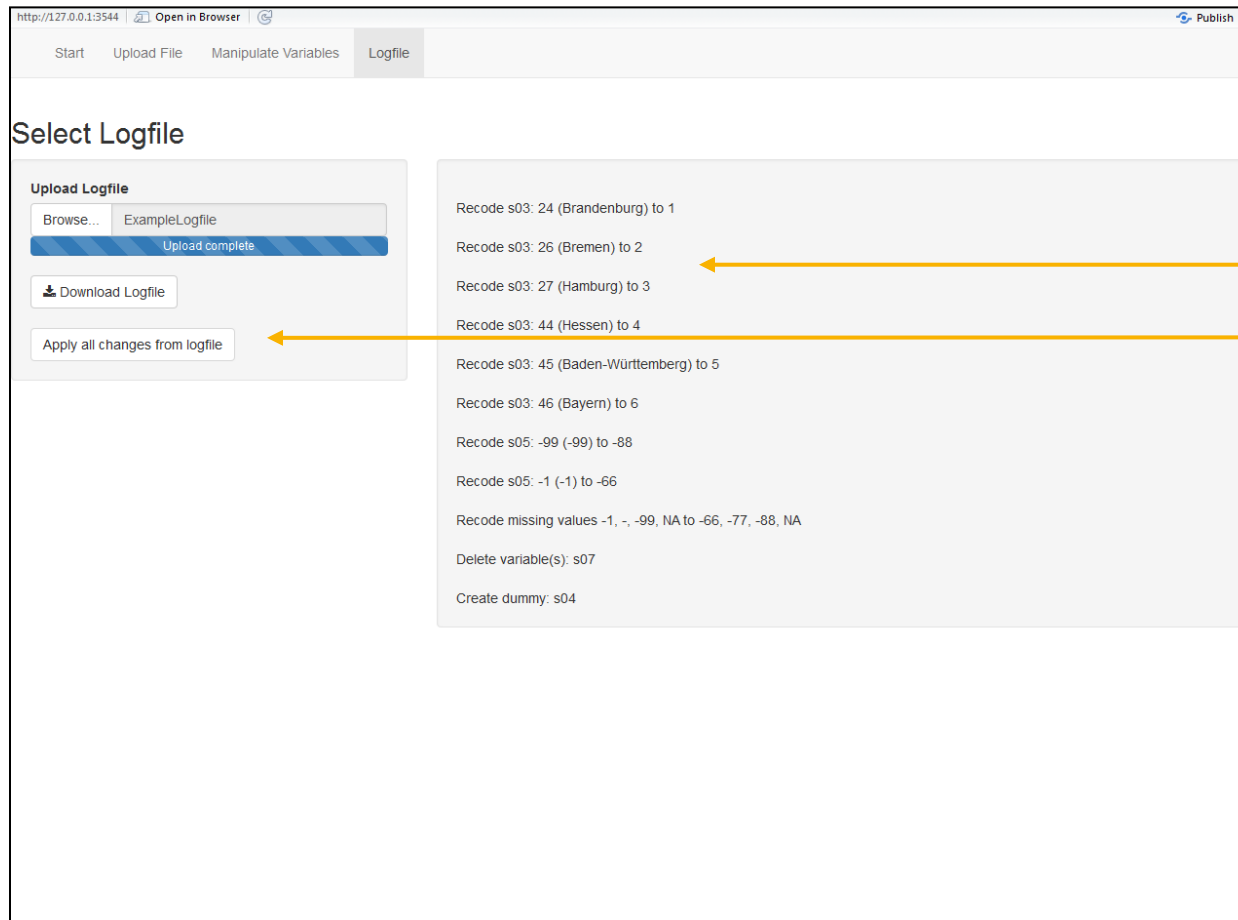
Browse... No file selected

Download Logfile

Upload an existing logfile here.

Or download the logfile of your current manipulations here.

Data Input



http://127.0.0.1:3544 Open in Browser Publish

Start Upload File Manipulate Variables Logfile

Select Logfile

Upload Logfile

Browse... ExampleLogfile

Upload complete

Download Logfile

Apply all changes from logfile

Recode s03: 24 (Brandenburg) to 1

Recode s03: 26 (Bremen) to 2

Recode s03: 27 (Hamburg) to 3

Recode s03: 44 (Hessen) to 4

Recode s03: 45 (Baden-Württemberg) to 5

Recode s03: 46 (Bayern) to 6

Recode s05: -99 (-99) to -88

Recode s05: -1 (-1) to -66

Recode missing values -1, -, -99, NA to -66, -77, -88, NA

Delete variable(s): s07

Create dummy: s04

After uploading a logfile, this gives you a summary of the changes that will be executed on the current dataset.

You can apply all these changes on your current dataset.

While all recodings will be done immediately, you can change the selection of variables that will be deleted or dummy coded under 'manipulate variables' later on.

Logfile

- The logfile is a simple csv-table.
- Each variable takes n rows, where n is the number of different levels.

variable	level	oldvalue	newvalue	MV1_old	MV1_new	MV2_old	MV2_new	MV3_old	MV3_new	MV4_old	MV4_new	Remain	Create Dummy	Priority
ID	1	1	1	NA	NA	NA	NA	NA	NA	NA	NA	1	0	1
ID	2	2	2	NA	NA	NA	NA	NA	NA	NA	NA	1	0	1
ID	3	3	3	NA	NA	NA	NA	NA	NA	NA	NA	1	0	1
ID	4	4	4	NA	NA	NA	NA	NA	NA	NA	NA	1	0	1
ID	5	5	5	NA	NA	NA	NA	NA	NA	NA	NA	1	0	1
ID	6	6	6	NA	NA	NA	NA	NA	NA	NA	NA	1	0	1
ID	7	7	7	NA	NA	NA	NA	NA	NA	NA	NA	1	0	1
ID	8	8	8	NA	NA	NA	NA	NA	NA	NA	NA	1	0	1
ID	9	9	9	NA	NA	NA	NA	NA	NA	NA	NA	1	0	1
ID	10	10	10	NA	NA	NA	NA	NA	NA	NA	NA	1	0	1
ID	11	11	11	NA	NA	NA	NA	NA	NA	NA	NA	1	0	1
ID	12	12	12	NA	NA	NA	NA	NA	NA	NA	NA	1	0	1
s03	Brandenburg	24	1	NA	NA	NA	NA	NA	NA	NA	NA	1	0	1
s03	Bremen	26	2	NA	NA	NA	NA	NA	NA	NA	NA	1	0	1
s03	Hamburg	27	3	NA	NA	NA	NA	NA	NA	NA	NA	1	0	1
s03	Hessen	44	4	NA	NA	NA	NA	NA	NA	NA	NA	1	0	1
s03	Baden-													
s03	Württemberg	45	5	NA	NA	NA	NA	NA	NA	NA	NA	1	0	1
s03	Bayern	46	6	NA	NA	NA	NA	NA	NA	NA	NA	1	0	1
s03	Berlin	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1	0	1
s04	1	1	1	NA	NA	NA	NA	NA	NA	NA	NA	1	1	3
s04	2	2	2	NA	NA	NA	NA	NA	NA	NA	NA	1	1	3
s04	3	3	3	NA	NA	NA	NA	NA	NA	NA	NA	1	1	3
s05	-99	-99	-88	-1	-66	-	-77	-99	-88	NA	NA	1	0	1
s05	-1	-1	-66	NA	NA	NA	NA	NA	NA	NA	NA	1	0	1

It seems like you can delete, dummify or set the priority of single levels of a variable, but you can only define it for the whole variable. However, here, only the first row of a variable will be considered by the app.

Missing values are saved in the first row of a variable. The following rows are meaningless.

