

YAC Text Recoder
version 2.07.a

User Guide

Table of Contents

| | |
|---|-----------|
| | 0 |
| Chapter I Introduction | 3 |
| 1 Requirements | 3 |
| 2 Installation | 3 |
| 3 Contact and Support | 3 |
| Chapter II Saving Recodes to the Data File | 5 |
| Chapter III Glossary | 7 |
| Chapter IV Demonstration Projects | 10 |
| Chapter V Keyboard Shortcuts | 12 |
| Chapter VI Ordering the Full Version | 14 |
| Chapter VII Changes | 16 |
| Chapter VIII Dialog Windows | 19 |
| 1 Auto Coding | 19 |
| 2 Edit Preferences... .. | 19 |
| Sets | 20 |

Chapter

」

1 Introduction

YAC Text Recoder is used to recode text responses (to open ended questions) into numeric data.

1.1 Requirements

YAC Text Recoder works on PC computers running one of the Microsoft Windows operating systems (Windows 9x, Windows NT, Windows 2000, Windows XP, and newer versions).

1.2 Installation

YAC Text Recoder needs no special installation - the packed file should be unpacked to any folder on the computer keeping the structure of subdirectories and locations of files in program's folders.

1.3 Contact and Support

In case of questions concerning the program or procedures described in this document, please contact:

YAC Software

contact@yac.com.pl

<http://www.yac.com.pl/>

Chapter

」

2 Saving Recodes to the Data File

When recoding responses in the recoding window, only the mapping response to code is saved in the project file. To save assigned codes as new columns / variables in the data file, you must run one of the commands:

Recode | Recode set or **Recode | Recode all**.

The first menu saves the data from the set currently being recoded. The second menu - from all sets.

Note

When saving codes to the data file, a column / variable is created for each question and each code. Then, a 1 is written in the appropriate cell based on the response to code mapping. Thus, these new columns may be analyzed as multi-response columns.

Note

When saving codes to the data file, the program writes data for existing response to code mapping. This means that if codes were written to the data file, then the mapping was changed (for instance, by joining two codes into one), and we try to save the codes again, all codes already written will remain in the data (these will not be cleared). However, new mappings will be added to the data file.

If you need to clear those old codes, the easiest approach is to restore the data file to its original version (before any codes were saved there) and run the save command again. Then only codes coming from the most recent response to code mapping will be saved to the data file.

Chapter



3 Glossary

Code

Codes group responses with the same meaning. For instance, for responses

apple
apples

you may want to define a code with the description of "apple", and for responses

orange
oranges

you may want to define another code with the description of "orange".

When defining a code, you also define the code's identifier - this identifier is displayed on the screen next to responses that have this code assigned to them. But more importantly, these codes are used to generate column / variable names when saving recodes to the data file. For dichotomous variables, the name of the variable has the format:

QuestionID "_" CodeID

So, for instance, for a question with the identifier "Q1" and codes with identifiers "A" and "B", the following two variables would be generated:

Q1_A
Q1_B

Column / variable

Columns / variables are the smallest elements in a data file that give separate responses, that is, responses from two columns can be recoded separately, but you cannot split a response that is written on a single column / variable (but you can still assign multiple codes to such responses).

Data file

This is the file from which the response texts are read from and where you will save all recodes after you are finished with recoding.

Currently, two formats are supported:

- SPSS data file (.sav),
- Fixed-ASCII data file (.asc or .dat).

Description file

For non-SPSS data files, describes the names and positions of columns / variables in the data file.

Currently, only the SPSS Syntax (.sps) format is supported for description files.

Question

Questions are groups of columns / variables that, when assigning codes to responses from these columns, treat all columns as being a single column with many possible responses.

For instance, assume that we have a spontaneous awareness question with three text variables: T1, T2, and T3. Next, let's define a question for these variables: Q1. Now, assume that we have the following responses in the data file:

| | T1 | T2 | T3 |
|---------|--------|--------|--------|
| case 1: | brandA | | |
| case 2: | brandC | | |
| case 3: | brandA | brandB | brandC |

Now, if we assign the codes A, B, and C to the respective responses, we will get the following numeric data:

| | Q1_A | Q1_B | Q1_C |
|---------|------|------|------|
| case 1: | 1 | 0 | 0 |
| case 2: | 0 | 0 | 1 |
| case 3: | 1 | 1 | 1 |

Project

A project consists of the following definitions:

- data file and description file,
- questions and columns / variables assigned to questions,
- sets and questions assigned to sets, as well as:
 - codes defined for all questions in a set,
 - recodes defined for responses to all questions in a set.

Project file

Projects are saved to project files with the `.trp` extension.

Recode

A recode in our application is the association of a code with a response.

Recodes are automatically saved to the project file. When you are finished with recoding and you want to analyze the recoded data, you should save those recodes to the data file (this will create new columns / variables for codes).

Response

A value of a column / variable for a single case in the data file. Usually, in the program, we will only talk about text (open) responses.

Set

Sets group questions that have similar responses (usually, questions in a set ask about the same thing for various entities or situations). When recoding a set, you don't need to handle each question separately, but you can use the same codes and recodes for all questions in a set.

For instance, the following questions would make good candidates for forming a single set:

- Why do you like beer A?
- Why do you like beer B?

But the following questions shouldn't be put into a single set:

- Why do you like summer?
- Where do you drink beer?

Chapter

IV

4 Demonstration Projects

The program is being distributed with two demonstration projects:

- in Polish (files `demo-pl.sav` and `demo-pl.trp`),
- in English (files `demo-en.sav` and `demo-en.trp`).

`.sav` files are SPSS data files that include responses to two open ended questions (`q1`, `q2`, and `q3`). In the first two questions, text responses for beer brand names were saved (and should be recoded as a set), the third question is about the beer's country of origin.

Users can either define [questions](#) and [sets](#) themselves (select the menu **Project | New** and then open one of the `.sav` files) or open an existing project file with definitions already in place (select the menu **Project | Open** and then open one of the `.trp` files).

Chapter

y

5 Keyboard Shortcuts

Most functions in the program can be accomplished (and accelerated) by using keyboard shortcuts:

| | |
|---|----------------------------|
| add question, set, code | Insert |
| edit question, set, code | F2 |
| delete question, set, code | Delete |
| open set | Enter |
| move from the response list to the codes list | Enter in the response list |
| add highlighted code to the response's recodes | Enter in the code list |
| add last used code to the current response's recodes and move to the next response | Gray Plus |

Chapter

VI

6 Ordering the Full Version

The following information concerns users of the demonstration version only.

The demonstration version of this application consists of the full-featured product that does not save project files, listings, nor the final numeric values to the data file. However, you can test all other functionality that is in the application.

If you need any information on purchasing the full version, prices and licensing, please contact:

YAC Software

contact@yac.com.pl

<http://www.yac.com.pl/>

Chapter

VII

7 Changes

Changes in consecutive versions of YAC Text Recoder (previously Text-Recoder) are documented below, starting with version 2.00.

Version 2.07.a

- Added handling of code descriptions containing line breaks - previously, defining codes with descriptions containing line breaks would break the format of the .trp project file and the program would no longer be able to open it.
- Fixed handling of long text variables in SPSS data files - after saving recodes to an SPSS data file, some long text variables would be visible in SPSS as a set of variables no longer than 255 characters.
- Fixed handling of SPSS data files with numerical values outside of the integer range (such as 1E20) - an "invalid floating point operation" error would be reported when trying to save recodes to the SPSS data file.

Version 2.07

- Added sorting by string similarity (in both the response list and the code list).

Version 2.06

- Added [automated coding](#).

Versions 2.05.a - 2.05.b

- A note on [saving codes to the data file](#) was added to this document.
- Various small fixes.

Version 2.05

- When importing data, the project file is checked for consistency.
- When importing data / opening Fixed-ASCII data, the data file is checked for consistency.
- Various small fixes.

Version 2.04

- Added handling of long text variables in SPSS data files (up to 32767 characters).
- Added a view box in the recode window that displays full text responses.

Version 2.03

- Added unification options for response texts:
 - Default options may be defined in the [preferences dialog window](#).
 - Options for sets may be defined in the set edit dialog window.

Version 2.02.a

- Fixed handling of SPSS data files with missing value ranges defined.

Version 2.02

- Faster opening and closing of data files.
- Faster reading of responses.
- Faster recoding of sets.
- Added progress bars when reading responses and recoding sets.

Versions 2.01.a - 2.01.c

- Fixed handling of SPSS long variable names.
- Importing old Text-Recoder projects enhanced.
- Various GUI fixes.

Version 2.01

- Added [demonstration data](#).

Version 2.00

- Multiple codes may be assigned to a single text response.
- Non-numeric code identifiers are supported.
- Direct access to SPSS data files (`.sav`) is supported.
- Code counts are reported on line in the code list of the recoding window.
- Recodes are written to the project file independently of saving numeric data to the data file.

Chapter

VIII

8 Dialog Windows

8.1 Auto Coding

This procedure automatically generates codes and assigns those codes to responses.

To run the procedure, select the **Auto coding** function from the **Tools** menu. A new dialog window will open, with the following options:

The **Codes** group - here, you can define how codes should be managed:

- **Use existing** - the algorithm will first assign responses to existing codes.
- **Define new** - if there are responses not assigned to any codes yet, the program will try to create new codes.
- **Clear existing** - all existing codes will be removed before starting the automatic coding procedure.

Minimum group count - this value will be used to stop the algorithm if newly created recodes have a lower response count than the given value.

Minimum response length - only responses of the given length, or longer, will be used in the procedure.

Note that short responses (1-2 characters long) are "similar" to almost any other text. Usually, you'll want to set this parameter to at least 3 or 4.

Algorithm

The auto coding procedure works as follows:

1)

All codes are removed, provided that the **Clear existing** option is checked.

2)

Existing codes will be used to create recodes for responses not recoded yet (provided that the **Use existing** option is checked).

All responses that weren't recoded yet are compared to the text of each existing code. If the response matches the text of the code, that code is assigned to the response.

3)

Creation of new codes.

From among all responses that are not recoded yet, the response with the highest count is taken. Next, all other non-recoded responses are compared with this response. If the sum of counts of all these matching responses (including the original response) is greater than or equal to the value of the **Minimum group count** parameter, a new recode is created.

The algorithm repeats this step until:

- either all responses have been recoded,
- or the size of the new recode is smaller than the **Minimum group count** parameter.

8.2 Edit | Preferences...

Several program-wide settings can be defined in this dialog window:

- language settings,
- extension associations.

8.2.1 Sets

Default text unification options can be defined in this tab:

- **Trim leading blanks** - deletes blanks from the front of the response,
- **Trim trailing blanks** - deletes blanks from the end of the response,
- **Remove repeating blanks** - contracts consecutive blanks into a single blank,
- **Unify case** - changes all texts to lowercase thus unifying responses that differ only by case,
- **Unify diacritics** - changes all diacritics into respective ANSI characters.

These options are automatically used when defining new sets.

Index

C

Contact 3

I

Installation 3

K

Keyboard shortcuts 12

R

Requirements 3

S

Support 3