



MyBox: Easy Tools Set

User Guide – Data Tools

Author: Mara

Version: 6.7

Date: 2023-1-21

Contents

1 Introduction.....	8
1.1 Main Interface.....	8
1.2 Resources Addresses.....	9
1.3 Documents.....	10
1.4 Menu of Tools.....	11
2 Data in Two-dimensional Storage Structure.....	12
2.1 Data Objects.....	12
2.2 Functions Menu.....	13
2.3 Define Columns	14
2.3.1 Interface of Columns Management.....	14
2.3.2 Types of Columns.....	15
2.3.3 Format of Column.....	16
2.3.3.1 Format for Numbers.....	16
2.3.3.2 Format for Datetime/Date.....	17
2.3.3.3 Format for Era.....	18
2.3.3.4 Define Enumeration.....	19
2.3.4 Handle Invalid Values.....	19
2.4 Usages of Columns.....	20
2.4.1 Load Data.....	20
2.4.2 Display Data.....	20
2.4.3 Controls for editing	21
2.4.4 Edit Data Cell.....	22
2.4.5 Save Data.....	24
2.4.6 Calculate Data.....	24
2.4.7 Sort Data.....	25
2.4.8 Color of Column.....	25
2.5 Edit Data.....	26
2.5.1 Mode "Table".....	26
2.5.2 Mode "CSV".....	27
2.5.3 Save and Recover.....	28
2.6 View Data.....	29
2.6.1 View Html.....	29
2.6.2 View CSV.....	30
2.7 Define Data Attributes.....	31
2.8 Row Expression.....	32
2.8.1 Usages of Row Expression.....	32
2.8.2 Edit Row Expression.....	32
2.8.3 Calculate Row Expression.....	33
2.8.4 Examples.....	33
2.9 Row Filter.....	34
2.9.1 Usages of Row Filter	34
2.9.2 Edit Row Filter.....	34
2.9.3 Exmaples.....	35

2.10 Group Rows.....	36
2.10.1 Before and after Grouping.....	36
2.10.2 Usages of grouping.....	36
2.10.3 Ways of Data Grouping.....	36
2.10.4 Group by Equal Values.....	37
2.10.4.1 Way of Grouping.....	37
2.10.4.2 Implementation of Grouping.....	37
2.10.4.3 Results of Grouping.....	38
2.10.5 Group by Value Ranges.....	39
2.10.5.1 Way of Grouping.....	39
2.10.5.2 Split Date/Era.....	40
2.10.5.3 Edit List of “Start-End”.....	41
2.10.5.4 Implementation of Grouping.....	42
2.10.5.5 Results of Grouping.....	42
2.10.6 Group by Time.....	43
2.10.6.1 Way of Grouping.....	43
2.10.6.2 Implementation of Grouping.....	43
2.10.6.3 Results of Grouping.....	44
2.10.7 Group by Expression.....	45
2.10.7.1 Way of Grouping.....	45
2.10.7.2 Implementation of Grouping.....	45
2.10.7.3 Results of Grouping.....	46
2.10.8 Group by Conditions.....	47
2.10.8.1 Way of Grouping.....	47
2.10.8.2 Implementation of Grouping.....	47
2.10.8.3 Results of Grouping.....	48
2.10.9 Group by Row Number.....	49
2.10.9.1 Way of Grouping.....	49
2.10.9.2 Implementation of Grouping.....	49
2.10.9.3 Results of Grouping.....	50
2.11 Modify Data.....	51
2.11.1 Add Rows.....	52
2.11.2 Set Values.....	53
2.11.3 Delete.....	54
2.11.4 Set styles / Mark Abnormal Values.....	55
2.11.4.1 Manage Styles.....	55
2.11.4.2 Data Scope.....	56
2.11.4.3 Define Styles.....	57
2.11.4.4 Apply Styles.....	58
2.11.5 Paste Content in System Clipboard.....	59
2.11.6 Paste Content in MyBox Clipboard.....	60
2.12 Trim Data.....	61
2.12.1 Copy/Filter/Query/Convert.....	62
2.12.2 Sort.....	63
2.12.3 Transpose	64
2.12.4 Normalization	65

2.12.5 Split/Group.....	66
2.13 Data Calculation.....	67
2.13.1 Row Expression.....	68
2.13.2 Descriptive Statistic	69
2.13.3 Statistic in Groups.....	70
2.13.3.1 Options of Calculation.....	70
2.13.3.2 Grouped Data.....	71
2.13.3.3 Statistic Data.....	71
2.13.3.4 Data in Chart.....	72
2.13.3.5 XY Chart.....	73
2.13.3.6 Pie Chart.....	74
2.13.4 Simple Linear Regression	75
2.13.4.1 Base of Implementation.....	75
2.13.4.2 Options of Calculation.....	75
2.13.4.3 Regression	75
2.13.4.4 Model.....	76
2.13.4.5 Fitted Chart.....	77
2.13.4.6 Residual Chart.....	78
2.13.5 Simple Linear Regression - Combination	79
2.13.6 Multiple Linear Regression.....	80
2.13.6.1 Regression.....	80
2.13.6.2 Model.....	81
2.13.7 Multiple Linear Regression - Combination.....	82
2.13.8 Frequency Distributions	83
2.13.9 Values Percentage	84
2.14 Data Charts.....	85
2.14.1 XY Chart.....	86
2.14.1.1 Data.....	86
2.14.1.2 Bar Chart.....	87
2.14.1.3 Stacked Bar Chart.....	88
2.14.1.4 Line Chart.....	88
2.14.1.5 Scatter Chart.....	89
2.14.1.6 Bubble Chart.....	89
2.14.1.7 Area Chart.....	90
2.14.1.8 Stacked Area Chart.....	91
2.14.1.9 Parameters of Data in Chart.....	91
2.14.1.10 Layout.....	92
2.14.1.11 Category Axis.....	93
2.14.1.12 Number Axis.....	94
2.14.2 Pie Chart.....	95
2.14.3 Box-and-whisker Chart	96
2.14.3.1 Way of Calculation	96
2.14.3.2 Options of Calculation	97
2.14.3.3 Statistic Data.....	98
2.14.3.4 Options of Chart.....	98
2.14.4 Self Comparison Bars Chart	99

2.14.4.1 Way of Calculation.....	99
2.14.4.2 Options of Calculation	100
2.14.4.3 Options of Chart.....	100
2.14.5 Comparison Bars Chart	102
2.14.5.1 Way of Calculation.....	102
2.14.5.2 Options of Calculation	102
2.14.5.3 Options of Chart.....	102
2.14.6 XYZ Chart.....	104
2.14.6.1 3-D Scatter.....	105
2.14.6.2 Surface Chart.....	106
2.14.7 Locations Distribution.....	108
2.14.7.1 Options of Data.....	108
2.14.7.2 Data in Map.....	108
2.14.7.3 Options of Map.....	109
2.14.7.4 Options of Playing.....	110
2.15 Group Charts.....	111
2.15.1 General Manufacture.....	112
2.15.1.1 Group Data.....	112
2.15.1.2 Results of Grouping.....	113
2.15.1.3 Display Options of Chart.....	114
2.15.1.4 Options of Playing.....	114
2.15.2 Group Data – XY Chart.....	115
2.15.2.1 Specific Options to Generate Chart.....	115
2.15.2.2 Specific Options to Display Chart.....	115
2.15.3 Group Data – Pie Chart.....	116
2.15.3.1 Specific Options to Generate Chart.....	116
2.15.4 Group Data – Box-and-whisker Chart.....	117
2.15.4.1 Specific Options to Generate Chart.....	117
2.15.4.2 Specific Options to Display Chart.....	117
2.15.5 Group Data – Self Comparison Bars Chart.....	118
2.15.5.1 Specific Options to Generate Chart.....	118
2.15.5.2 Specific Options to Display Chart.....	118
2.15.6 Group Data – Comparison Bars Chart.....	119
2.15.6.1 Specific Options to Generate Chart.....	119
2.15.6.2 Specific Options to Display Chart.....	119
2.16 Input Data.....	120
2.16.1 Load Contents in System Clipboard	120
2.16.2 Import Examples.....	121
2.16.2.1 Personal Data.....	121
2.16.2.2 Statistic Data of China.....	122
2.16.2.3 Data of Regression	123
2.16.3 Location Data.....	124
2.17 Output Data.....	125
2.17.1 Export.....	126
2.17.1.1 Data Source.....	126
2.17.1.2 Formats of Targets.....	127

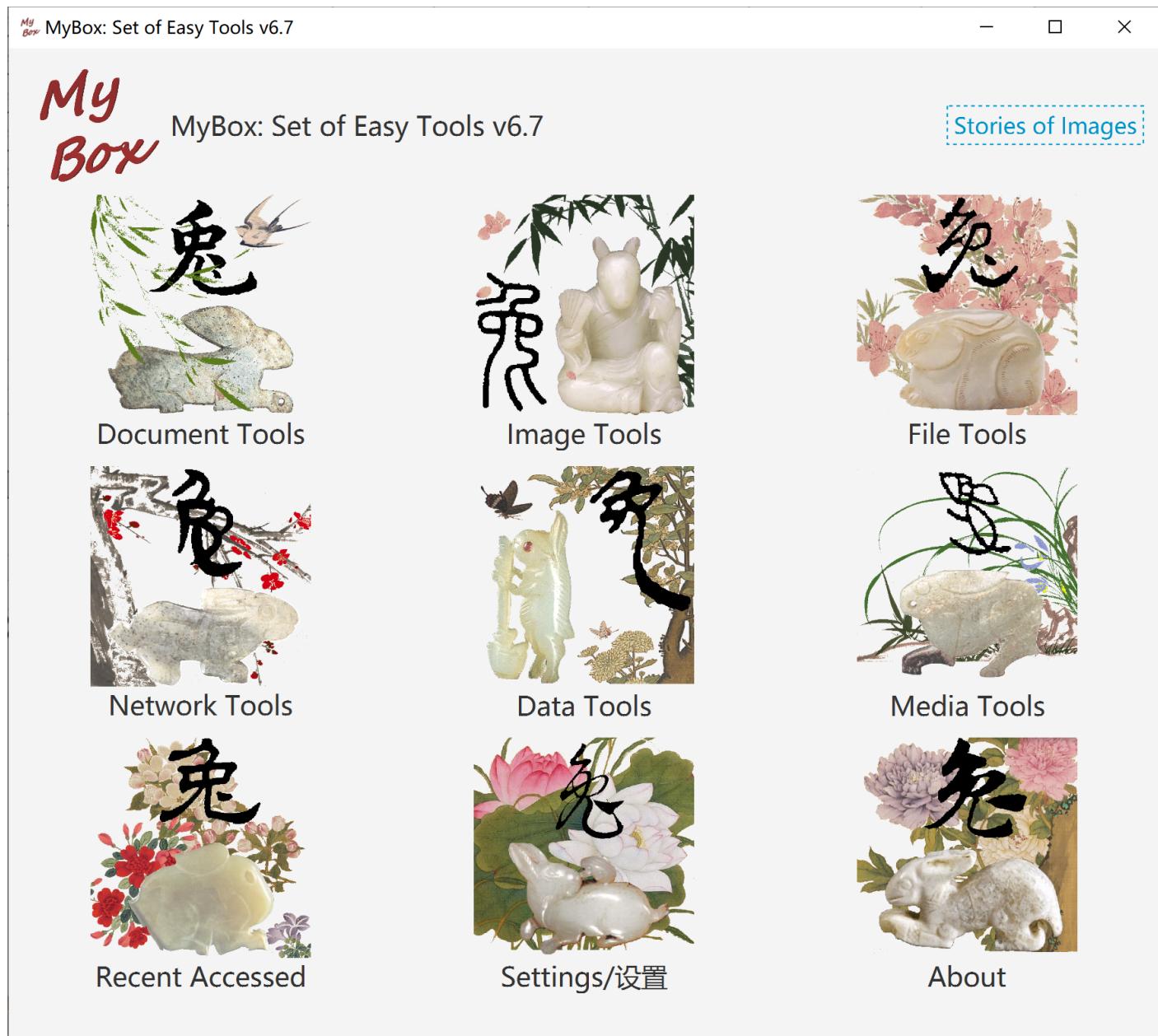
2.17.1.3 Target Files.....	127
2.17.2 Convert to Database Table.....	128
2.18 Manage Data.....	129
2.19 Splice Data.....	130
2.20 Data File.....	131
2.20.1 CSV File.....	131
2.20.2 Text File.....	132
2.20.3 Excel File.....	133
2.20.4 Convert/Split Data Files in Batch.....	134
2.20.5 Merge Data Files	135
2.21 Data in System Clipboard.....	136
2.22 Data in MyBox Clipboard.....	137
2.23 Matrix.....	138
2.23.1 Edit and Manage Matrices.....	138
2.23.2 Unary Matrix Calculation.....	139
2.23.3 Binary Matrices Calculation.....	140
2.24 Database Tables.....	141
2.24.1 Manage Database Tables.....	141
2.24.2 Limitations of SQL Identifier.....	142
2.24.3 Database SQL.....	143
2.24.4 SQL Query.....	144
3 Script and Expression.....	145
3.1 JShell(Java interactive coding tool).....	145
3.1.1 About JShell.....	145
3.1.2 Run Jshell in GUI.....	145
3.2 JEXL(Java Expression Language).....	147
3.2.1 About JEXL.....	147
3.2.2 Run JEXL in GUI.....	147
3.3 Javascript	149
4 Math Function.....	150
4.1 Define Math Functions.....	150
4.2 Calculate Math Function.....	150
4.3 Data Set.....	152
4.4 XY Chart of Unary Function.....	153
4.5 XYZ Chart of Binary Function.....	154
5 Data of Location.....	155
5.1 Data Constraints.....	155
5.1.1 Invalid Value.....	155
5.1.2 Coordinate System:.....	155
5.1.3 Coordinate Values.....	155
5.1.4 Time.....	155
5.1.4.1 Date Formats.....	155
5.1.4.2 Era	155
5.1.4.3 Examples	156
5.2 Data Operations.....	156
5.3 Map Data.....	157

5.4 Geography Code.....	159
5.4.1 Data Definition.....	159
5.4.2 Data Constraints.....	159
5.4.3 Edit Data.....	159
5.4.4 Define Condition.....	159
5.4.5 Import Data.....	159
5.4.5.1 Predefined Data.....	159
5.4.5.2 CSV Format.....	160
5.4.5.3 Data from geoname.org:.....	160
5.4.6 Settings	160
5.5 Location in Map.....	162
5.6 Location Tools.....	163
6 Others.....	164
6.1 Create Barcodes.....	164
6.2 Decode Barcodes.....	164
6.3 Message Digest.....	164
6.4 Encode/Decode Base64.....	164
6.5 Extract ttf files from ttc file.....	164

1 Introduction

This is desktop application based on JavaFx to provide simple and easy functions. It's free and open sources.

1.1 Main Interface



1.2 Resources Addresses

Contents	Link
Project Main Page	https://github.com/Mararsh/MyBox/
Source Codes and Compiled Packages	https://github.com/Mararsh/MyBox/releases
Submit Software Requirements and Problem Reports	https://github.com/Mararsh/MyBox/issues
Data	https://github.com/Mararsh/MyBox_data
Documents	https://github.com/Mararsh/MyBoxDoc
Mirror Site	https://sourceforge.net/projects/mara-mybox/files/
Cloud Storage	https://pan.baidu.com/s/1fWMRzym_jh075OCX0D8yA#list/path=%2F

The screenshot shows a GitHub repository page for 'Mararsh / MyBox'. The top navigation bar includes links for Pull requests, Issues, Marketplace, and Explore. Below the navigation bar, there's a search bar and a 'Code' button highlighted with a red circle. The main content area shows a list of files and their commit history. On the right side, there's an 'About' section with a list of tools and features, and a 'Releases' section with a recent release entry for 'v6.5.8'.

About

- Easy tools of document, image, file, network, location, color, and media.
- html markdown pdf image ocr
- csv sql database excel convert
- location javafx media bytes ppt
- jshell matrix

Releases 89

v6.5.8 Latest
18 days ago

+ 88 releases

Code

Releases

ReadMe in English

<https://github.com/Mararsh/MyBox/releases>

1.3 Documents

Name	Version	Time	English	Chinese
Development Logs	6.7	2023-1-21	html	html
Shortcuts	6.5.6	2022-6-11	html	html
Functions list	6.6.2	2022-11-30	html	html
Packing Steps	6.3.3	2020-9-27	html	html
Development Guide	2.1	2020-8-27	PDF	PDF
User Guide - Overview	6.7	2023-1-21	html PDF odt	html PDF odt
User Guide - Data Tools	6.7	2023-1-21	html PDF odt	html PDF odt
User Guide - Document Tools	6.7	2023-1-21	html PDF odt	html PDF odt
User Guide - Image Tools	6.7	2023-1-21	html PDF odt	html PDF odt
User Guide - File Tools	6.7	2023-1-21	html PDF odt	html PDF odt
User Guide - Network Tools	6.7	2023-1-21	html PDF odt	html PDF odt
User Guide - Media Tools	6.7	2023-1-21	html PDF odt	html PDF odt
User Guide - Development Tools	6.7	2023-1-21	html PDF odt	html PDF odt
Examples - Notes	6.6.1	2022-11-16	html	html
Examples - Information in Tree	6.6.1	2022-11-16	html	html
Examples - Favorite Address	6.6.1	2022-11-16	html	html
Examples - SQL	6.6.1	2022-11-16	html	html
Examples - JShell	6.6.1	2022-11-16	html	html
Examples - JEXL	6.6.1	2022-11-16	html	html
Examples - JavaScript	6.6.1	2022-11-16	html	html
Examples - Math Function	6.6.1	2022-11-16	html	html
Examples - Row Filter	6.6.1	2022-11-16	html	html
About - Color	6.6.1	2022-11-16	html	html
About - Coordinate System	6.6.1	2022-11-16	html	html
About - Media	6.6.1	2022-11-16	html	html
About - Data Analysis	6.6.1	2022-11-16	html	html
Palette - Common Web Colors	6.7	2023-1-21	major all	major all
Palette - Chinese Traditional Colors	6.7	2023-1-21	major all	major all
Palette - Japanese Traditional Colors	6.7	2023-1-21	major all	major all
Palette - Colors from colorhexa.com	6.7	2023-1-21	major all	major all
Palette - MyBox Colors	6.7	2023-1-21	major all	major all
Stories of Images	6.7	2023-1-21	html PDF odt	html PDF odt

1.4 Menu of Tools

MyBox: Set of Easy Tools v6.7

Stories of Images

The screenshot shows the MyBox application window with several tool icons:

- Document Tools:** Features a large stylized character '兔' (Rabbit) and a smaller rabbit illustration.
- Image:** Features a stylized character '兔' and a traditional Chinese figure.
- Network Tools:** Features a stylized character '兔' and a landscape illustration.
- Recent Access:** Features a stylized character '兔' and a floral illustration.

A context menu is open over the 'Image' icon, listing the following options:

- Manufacture Data
- Manage Data
- Splice Data
- Row filter
- Data File** (selected)
- Matrix
- Database
- Data in System Clipboard
- Data in MyBox Clipboard
- Script and Expression
- Math Function
- Location
- Miscellaneous
- Close(ESC/F6 Or click anywhere outside the object)**

The 'Edit CSV File' option under 'Data File' is also highlighted.

2 Data in Two-dimensional Storage Structure

2.1 Data Objects

1. Following objects can be edited in consistent way:

- Data files, including CSV File, Excel file, texts file.
- Data of MyBox Clipboard
- Matrices
- Database tables.

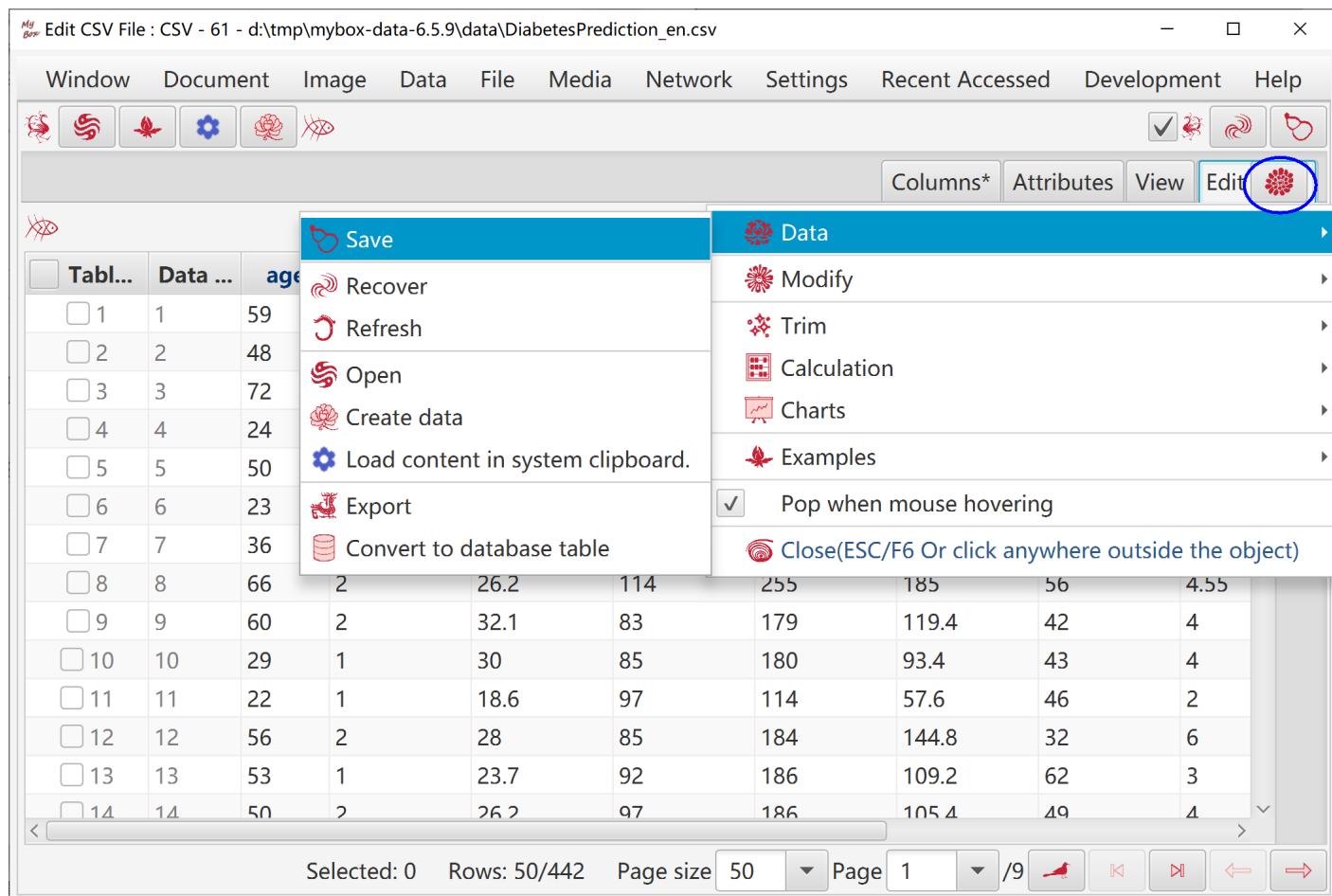
2. Data are represented as two-dimensional storage structure:

- "Columns" define dimensions of data in horizontal direction.
- "Rows" save values of data in vertical direction..
- Data should be in same width. That is all rows have equal number of columns.

Tabl...	Data ...	age					
1	1	59					
2	2	48					
3	3	72					
4	4	24					
5	5	50					
6	6	23					
7	7	36					
8	8	66					
9	9	60					
10	10	29	1	30	85	180	93.4
11	11	22	1	18.6	97	114	57.6
12	12	56	2	28	85	184	144.8
13	13	53	1	23.7	92	186	109.2
14	14	50	2	26.2	97	186	105.4

2.2 Functions Menu

Hover or click button “Functions” to pop Functions Menu.



2.3 Define Columns

2.3.1 Interface of Columns Management

Under tab “Columns”, add/delete/change columns in table view:

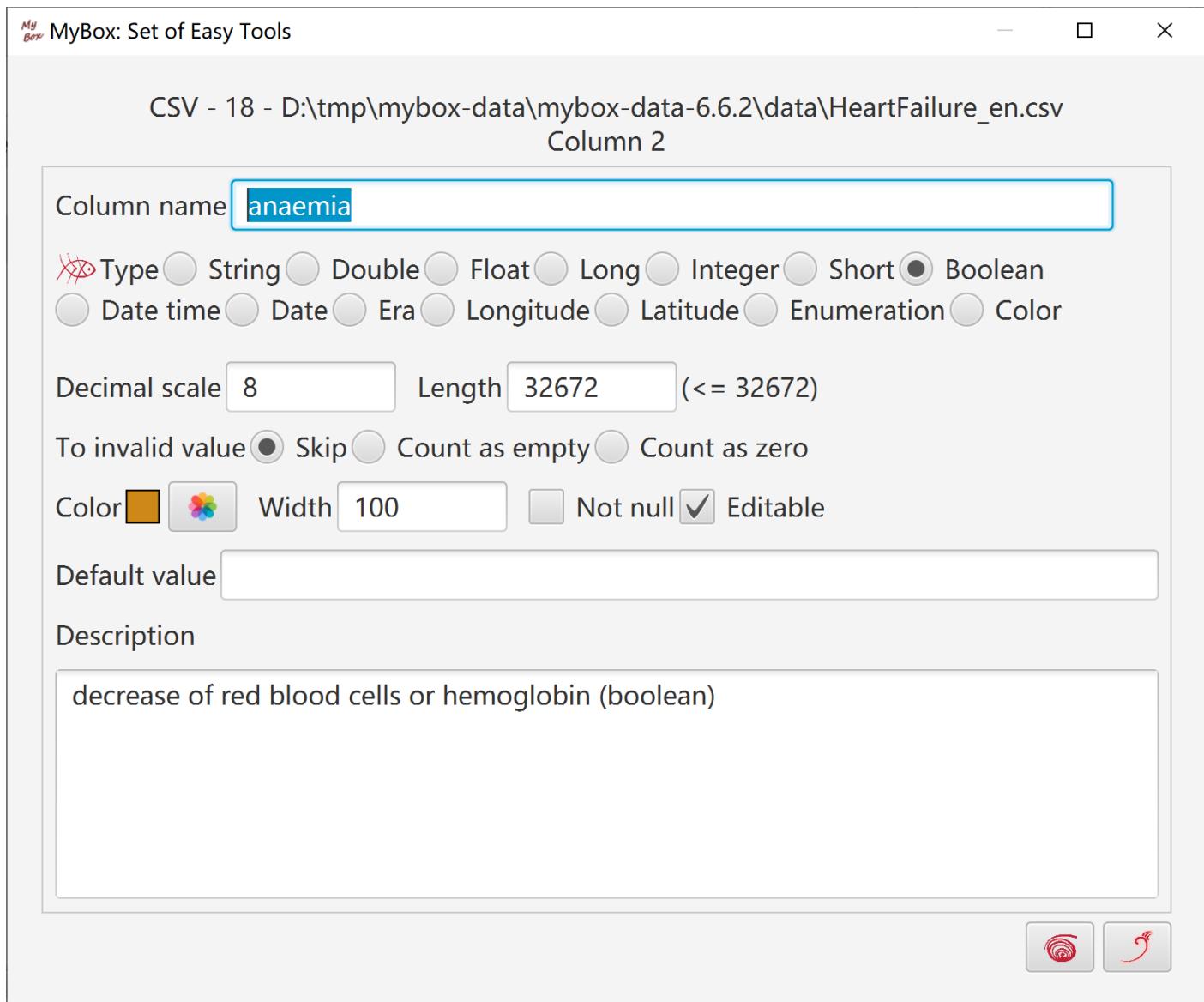
1. Column names should not be null nor duplicated.
2. Click table cell to edit value directly.
3. Select one row and click button “Edit” to open edit window.
4. Can rename all columns with sequence numbers.
5. Can set random colors.
6. Can adjust orders of columns.
7. Click button "OK" to apply modifications of columns to "Table" of current data.
8. Click button "Cancel" to discard modifications of columns and pick data from "Table" of current data.
9. When changes have not been saved, if modifications have not been applied to data, “**” is displayed in tab header, or else “*” is displayed in tab header.

Ta...	Index	Column name	Type	Color	Width	Editable	Not null	Separat...
<input type="checkbox"/>	1	age	Double	[Color Box]	61	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<input type="checkbox"/>	2	sex	Short	[Color Box]	100	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<input type="checkbox"/>	3	BMI(body mass index)	Double	[Color Box]	-----	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	4	BP(average blood pressure)	Double	[Color Box]	0x9E556AFF	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<input type="checkbox"/>	5	S1(blood serum measurement 1)	Double	[Color Box]	#9E556A	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<input type="checkbox"/>	6	S2(blood serum measurement 2)	Double	[Color Box]	-6400662	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<input type="checkbox"/>	7	S3(blood serum measurement 3)	Double	[Color Box]	sRGB: 158 85 106 100%	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<input type="checkbox"/>	8	S4(blood serum measurement 4)	Double	[Color Box]	HSB: 343 46% 62%	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<input type="checkbox"/>	9	S5(blood serum measurement 5)	Double	[Color Box]	Adobe RGB: 141 86 105	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<input type="checkbox"/>	10	S6(blood serum measurement 6)	Double	[Color Box]	Apple RGB: 137 64 87	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<input type="checkbox"/>	11	disease progression one year after ...	Double	[Color Box]	ECI RGB: 148 93 115	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Selected: 0 Rows: 50/442 Page size 50

2.3.2 Types of Columns

1. Types of columns include: String, Double, Float, Long, Integer, Short, Boolean, Datetime, Date, Era, Longitude, Latitude, Enumeration, Color.
2. This attribute is used to display, edit, calculate, and save data.
3. Longitude and Latitude are defined together generally.

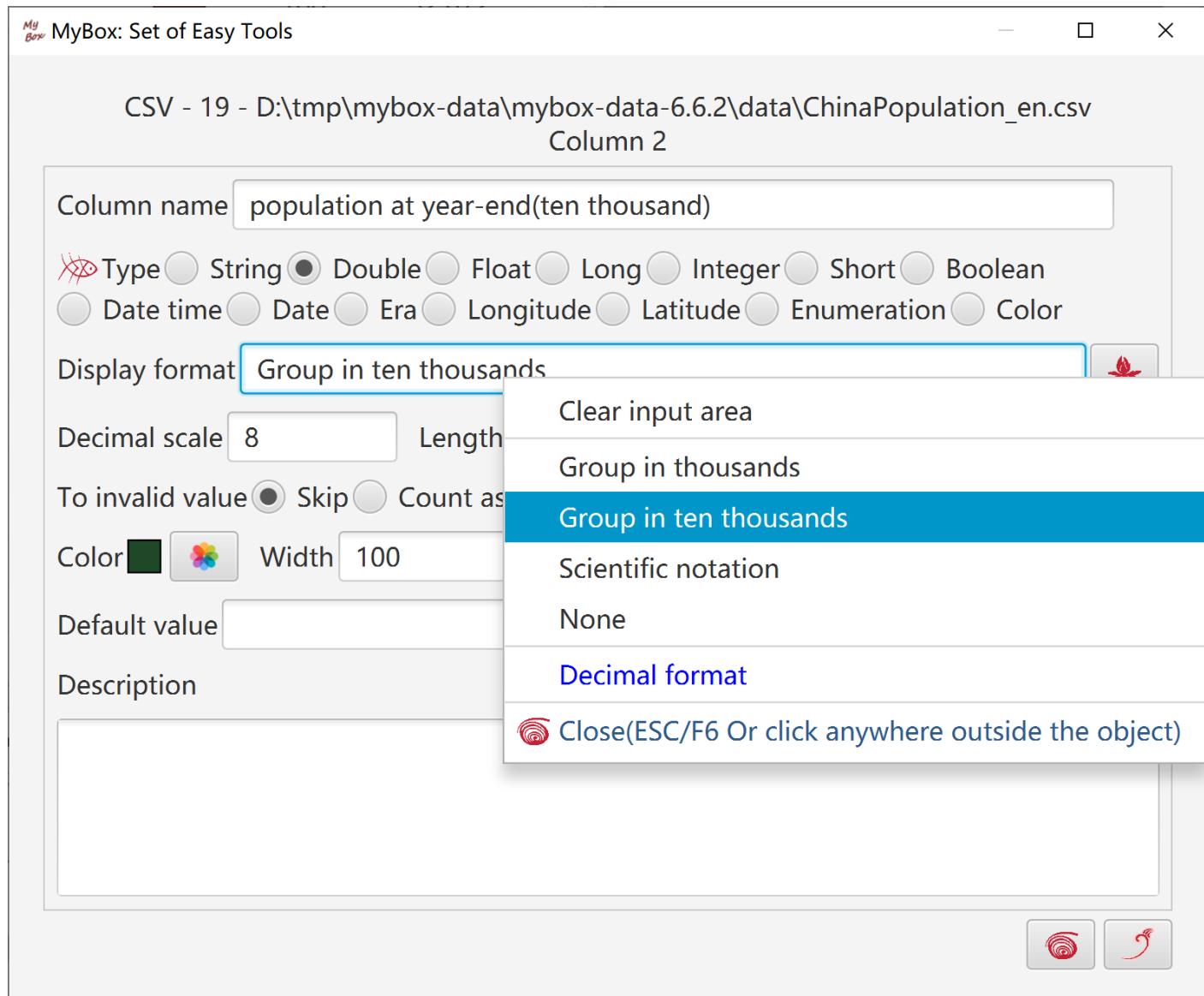


2.3.3 Format of Column

1. This attribute is mainly for display. When data are inputted/edited, formats are not applied automatically and original inputs are kept.
2. In some interfaces, like "Copy" or "Export", options "Save date/time/era and numbers as columns' formats" can be checked.

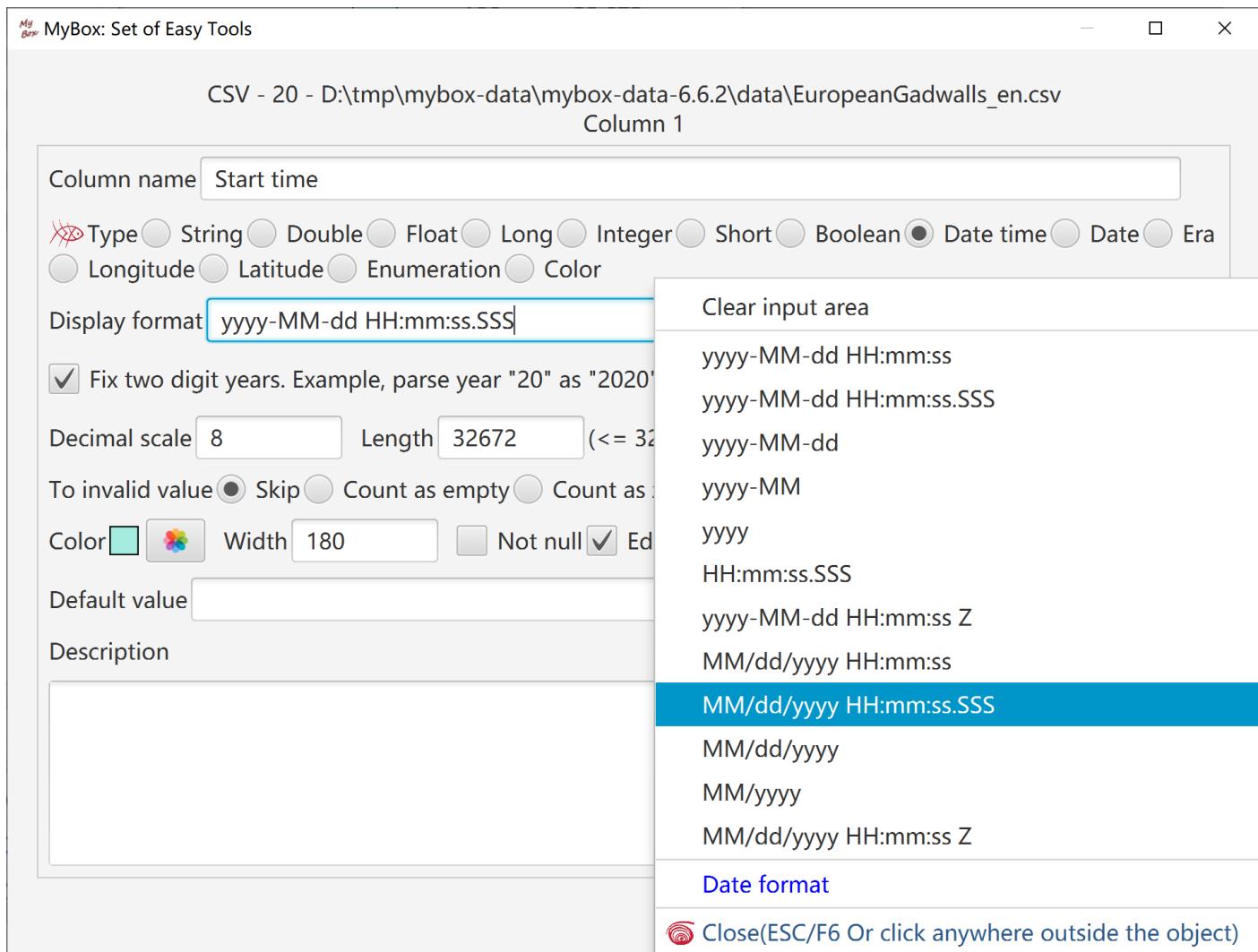
2.3.3.1 Format for Numbers

To numbers, format can be: group in thousands, group in ten thousands, scientific notation, and no format.



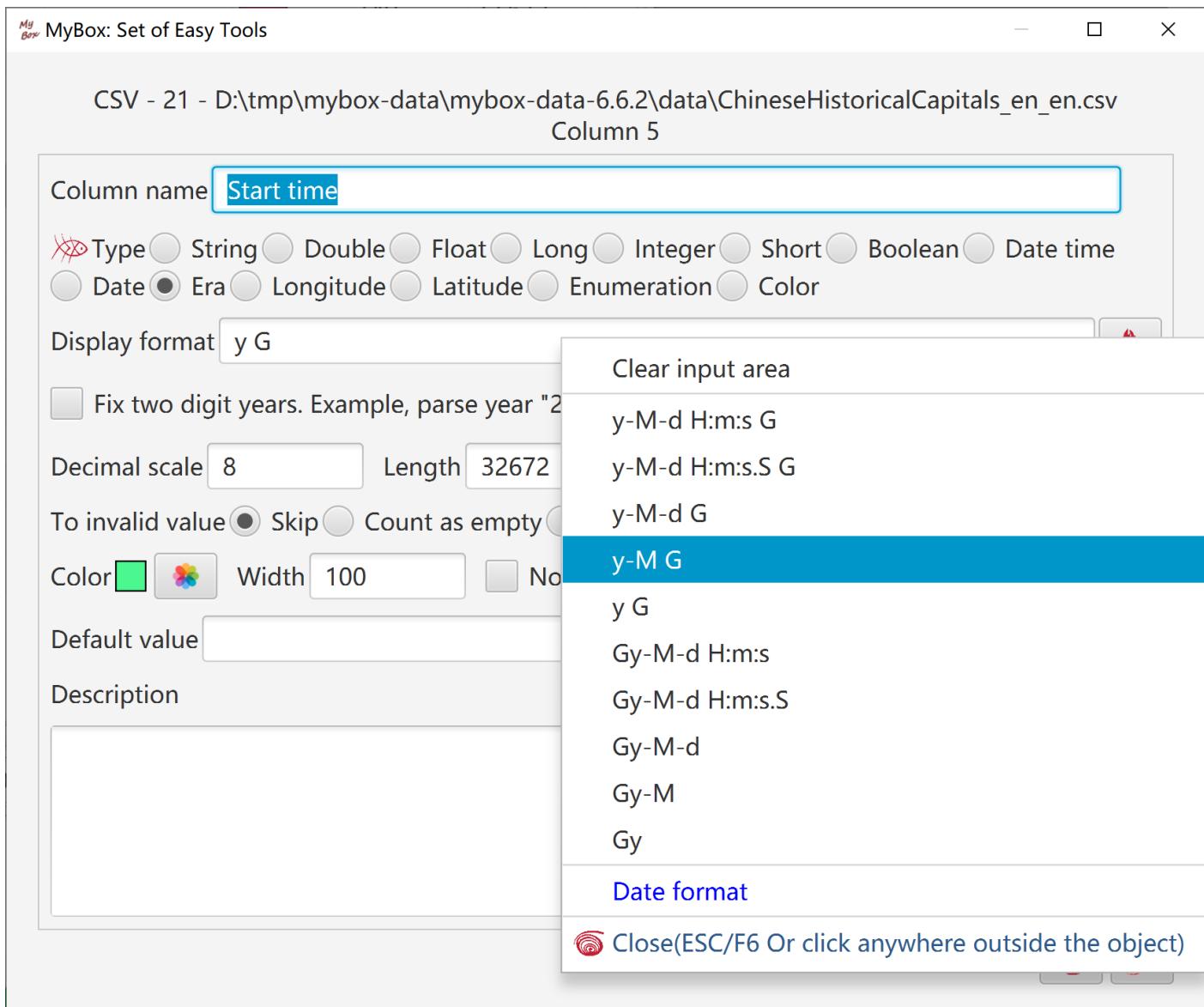
2.3.3.2 Format for Datetime/Date

To Datetime/Date, following are supported in formats: MM/dd/yy, yy-MM-dd, milliseconds, time zone, T separator, patch century, etc..



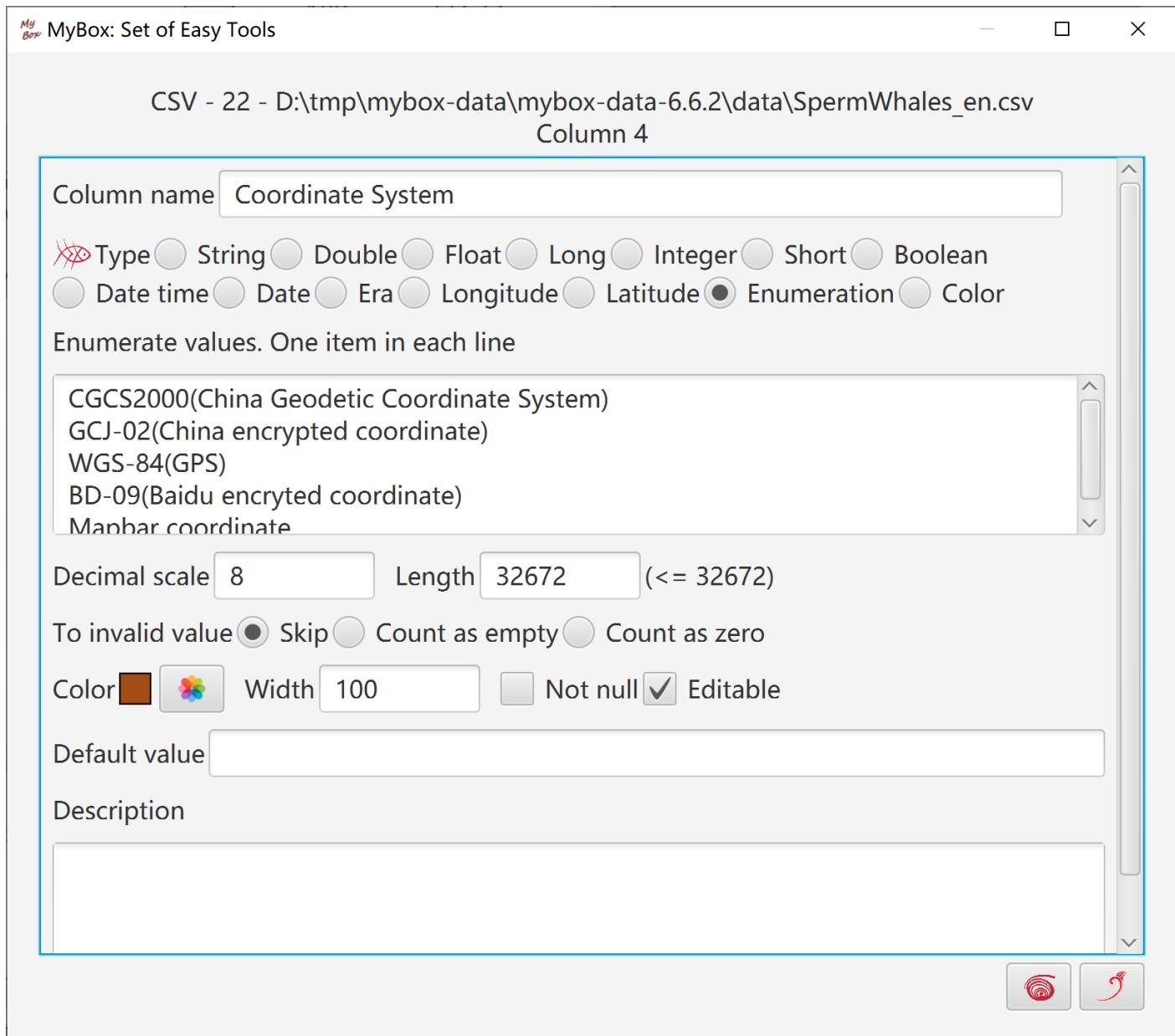
2.3.3.3 Format for Era

To Era, following are supported in formats: MM/dd/yy, yy-MM-dd, milliseconds, time zone, T separator, patch century, prefix/suffix of “AD” and “BC” in Chinese and English,etc..



2.3.3.4 Define Enumeration

To Enumeration, list of values can be defined.



2.3.4 Handle Invalid Values

1. How columns handle invalid values, including: skip, count as empty, and count as zero.
2. In some context, “count as empty” equals to “skip”.
3. This attribute is only used for display or calculation. When data are inputted/edited, invalid values are not handled automatically.

2.4 Usages of Columns

Principle of column usages is “Most tolerability and least manufacture”.

2.4.1 Load Data

When load data, types of columns are not checked, and original values are read and imported.

2.4.2 Display Data

1. Parse values as columns' types.
2. Handle invalid values as columns' definitions.
3. Rewrite values as columns' formats.
4. Displayed values may be different from current actual values.
5. Example, Column type is Double and value “abc” is read:
 - If count invalid as zero, then display it as zero(actual value is still “abc”).
 - If skip or count invalid as empty, then display it as “abc”.

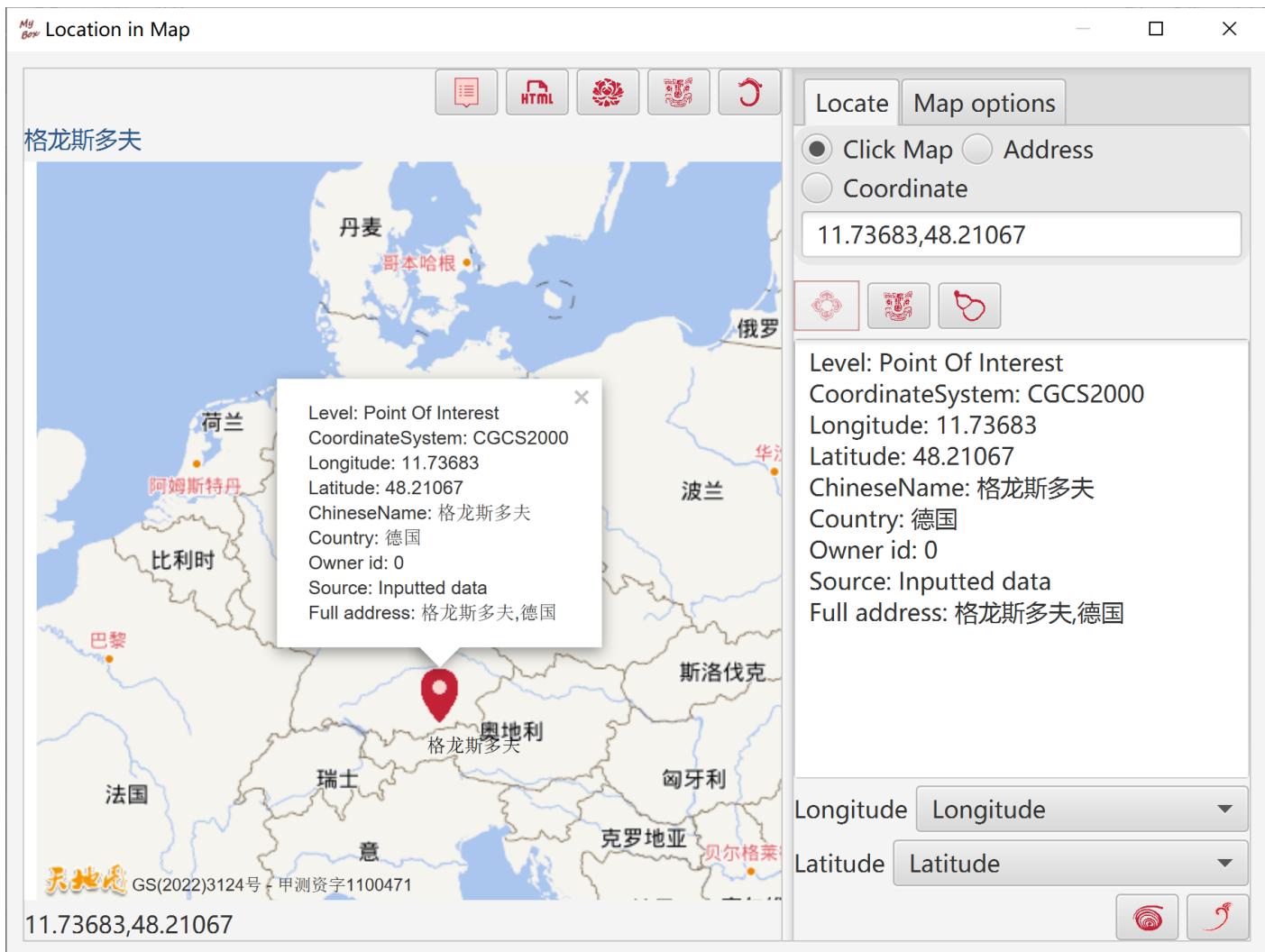
Table r...	Data...	year_	populatio...	male(ten t...	female(te...	urban(ten ...	rural(ten t...	
<input type="checkbox"/>	1	1949	5,4167	28,145	26022	5765	48402	
<input type="checkbox"/>	2	1950	5,5196	28,669	26527	6169	49027	
<input type="checkbox"/>	3	1951	5,6300	29,231	27069	6632	49668	
<input type="checkbox"/>	4	1952	5,7482	29,833	27649	7163	50319	
<input checked="" type="checkbox"/>	5	1953	58796	30,468	28328	7826	50970	
<input type="checkbox"/>	6	1954	6,0266	31,242	29024	8249	52017	
<input type="checkbox"/>	7	1955	6,1465	31,809	29656	8285	53180	
<input type="checkbox"/>	8	1956	6,2828	32,536	30292	9185	53643	
<input type="checkbox"/>	9	1957	6,4653	33,469	31184	9949	54704	
<input type="checkbox"/>	10	1958	6,5994	34,195	31799	10721	55273	
<input type="checkbox"/>	11	1959	6,7207	34,890	32317	12371	54836	
<input type="checkbox"/>	12	1960	6,6207	34,283	31924	13073	53134	
<input type="checkbox"/>	13	1961	6,5859	33,880	31979	12707	53152	
<input type="checkbox"/>	14	1962	6,7296	34,517	32778	11659	55636	

2.4.3 Controls for editing

- To Boolean, checkbox is provided.
- To Enumeration, list view is provided with selections.
- To Color, palette is provided.
- To Longitude and Latitude, map can be popped to locate coordinate.

Edit CSV File : CSV - 943 - d:\tmp\mybox-data-6.5.9\data\HeartFailure_en.csv *

Table r...	Data...	age	anaemia	creatinine...	diabetes	ejection fr...	high bloo...	platelets(k...	
<input type="checkbox"/> 1	1	61-80	<input type="checkbox"/>	582	<input type="checkbox"/>	20	<input checked="" type="checkbox"/>	265000	1
<input type="checkbox"/> 2	2	41-60	<input type="checkbox"/>	7861	<input type="checkbox"/>	38	<input type="checkbox"/>	263358.03	1
<input type="checkbox"/> 3	3	61-80	<input type="checkbox"/>	146	<input type="checkbox"/>	20	<input type="checkbox"/>	162000	1
<input type="checkbox"/> 4	4	41-60	<input checked="" type="checkbox"/>	111	<input type="checkbox"/>	20	<input type="checkbox"/>	210000	1
<input type="checkbox"/> 5	5	61-80	<input checked="" type="checkbox"/>	160	<input checked="" type="checkbox"/>	20	<input type="checkbox"/>	327000	2
<input checked="" type="checkbox"/> 6	6	81-...	<input checked="" type="checkbox"/>	47	<input type="checkbox"/>	40	<input checked="" type="checkbox"/>	204000	2
<input type="checkbox"/> 7	7	1-20	<input checked="" type="checkbox"/>	246	<input type="checkbox"/>	15	<input type="checkbox"/>	127000	1
<input type="checkbox"/> 8	8	21-40	<input checked="" type="checkbox"/>	315	<input checked="" type="checkbox"/>	60	<input type="checkbox"/>	454000	1
<input type="checkbox"/> 9	9	41-60	<input type="checkbox"/>	157	<input type="checkbox"/>	65	<input type="checkbox"/>	263358.03	1
<input type="checkbox"/> 10	10	61-80	<input checked="" type="checkbox"/>	123	<input type="checkbox"/>	35	<input checked="" type="checkbox"/>	388000	9
<input type="checkbox"/> 11	11	81-100	<input checked="" type="checkbox"/>	81	<input type="checkbox"/>	38	<input checked="" type="checkbox"/>	368000	4
<input type="checkbox"/> 12	12	61-80	<input type="checkbox"/>	231	<input type="checkbox"/>	25	<input checked="" type="checkbox"/>	253000	0
<input type="checkbox"/> 13	13	41-60	<input checked="" type="checkbox"/>	981	<input type="checkbox"/>	30	<input type="checkbox"/>	136000	1
<input type="checkbox"/> 14	14	41-60	<input checked="" type="checkbox"/>	168	<input type="checkbox"/>	38	<input checked="" type="checkbox"/>	276000	1



2.4.4 Edit Data Cell

- When user clicks a data cell in table view, editing is started. Its original value is displayed, while both type and format of column are ignored.
- While user inputs and modifies the value, value in the edit control is checked by column type:
 - If value is invalid, then edit control is displayed in abnormal color.
 - If value is valid, then edit control is displayed in normal color.
 - The value is always kept as what user has inputted.
- When focus leaves the data cell, value in the edit control is checked:
 - If value is not changed, then no checking of column type nor saving.
 - If value is changed, then check the changed value as column type:
 - To invalid value, discard it, and then restore and display original value.
 - To valid value, submit and save as new value, and then display saved value as type and format of the column.

4. Example, column type is Double, count invalid as empty, and decimal scale is 2. When read “abc”:
 - The data cell is displayed as “abc”.
 - User modifies it as “abc123”:
 - While user inputs the change, the text field is always in abnormal color.
 - After user clicks Enter or other control, this data cell comes back as “abc” in normal color.
 - User modifies it as “123.4567”:
 - While user inputs the change, the text field is always in normal color.
 - After user clicks Enter or other control, this data cell is saved as “123.4567” and displayed as “123.46”.
5. Other data cells are not affected. That is, data cells are always in original values if they are not changed.

MyBox Edit CSV File : CSV - 943 - d:\tmp\mybox-data-6.5.9\data\HeartFailure_en.csv

Window Document Image Data File Media Network Settings Recent Accessed Development Help

Columns Attributes View Edit

Table CSV

Selected: 1 Rows: [1-50]50/299 Page size 50 Page 1 /6

Table r...	Data...	age	anaemia	creatinine...	diabetes	ejection fr...	high bloo...	platelets(k...	...
<input type="checkbox"/> 1	1	75	<input type="checkbox"/>	582	<input type="checkbox"/>	20	<input checked="" type="checkbox"/>	265000	1
<input type="checkbox"/> 2	2	55	<input type="checkbox"/>	7861	<input type="checkbox"/>	38	<input type="checkbox"/>	263358.03	1
<input type="checkbox"/> 3	3	65	<input type="checkbox"/>	146	<input type="checkbox"/>	20	<input type="checkbox"/>	162000	1
<input type="checkbox"/> 4	4	50	<input checked="" type="checkbox"/>	111	<input type="checkbox"/>	20	<input type="checkbox"/>	210000	1
<input type="checkbox"/> 5	5	65	<input checked="" type="checkbox"/>	160	<input checked="" type="checkbox"/>	20	<input type="checkbox"/>	327000	2
<input type="checkbox"/> 6	6	90	<input checked="" type="checkbox"/>	47	<input type="checkbox"/>	40	<input checked="" type="checkbox"/>	204000	2
<input checked="" type="checkbox"/> 7	7	75	<input checked="" type="checkbox"/>	246abd	<input type="checkbox"/>	15	<input type="checkbox"/>	127000	1
<input type="checkbox"/> 8	8	60	<input checked="" type="checkbox"/>	315	<input checked="" type="checkbox"/>	60	<input type="checkbox"/>	454000	1
<input type="checkbox"/> 9	9	65	<input type="checkbox"/>	157	<input type="checkbox"/>	65	<input type="checkbox"/>	263358.03	1
<input type="checkbox"/> 10	10	80	<input checked="" type="checkbox"/>	123	<input type="checkbox"/>	35	<input checked="" type="checkbox"/>	388000	9
<input type="checkbox"/> 11	11	75	<input checked="" type="checkbox"/>	81	<input type="checkbox"/>	38	<input checked="" type="checkbox"/>	368000	4
<input type="checkbox"/> 12	12	62	<input type="checkbox"/>	231	<input type="checkbox"/>	25	<input checked="" type="checkbox"/>	253000	0
<input type="checkbox"/> 13	13	45	<input checked="" type="checkbox"/>	981	<input type="checkbox"/>	30	<input type="checkbox"/>	136000	1
<input type="checkbox"/> 14	14	50	<input checked="" type="checkbox"/>	168	<input type="checkbox"/>	38	<input checked="" type="checkbox"/>	276000	1
<input type="checkbox"/> 15	15	49	<input checked="" type="checkbox"/>	80	<input type="checkbox"/>	30	<input checked="" type="checkbox"/>	427000	1
<input type="checkbox"/> 16	16	82	<input checked="" type="checkbox"/>	379	<input type="checkbox"/>	50	<input type="checkbox"/>	47000	1

2.4.5 Save Data

1. Values are written into file of CSV/Texts/Excel as strings.
2. Values are written into database table as nearest types:

Column Type of MyBox	Data Type of JDBC
String	VARCHAR
Double	DOUBLE
Float	FLOAT
Long	BIGINT
Integer	INT
Short	SMALLINT
Boolean	BOOLEAN
Datetime	TIMESTAMP
Date	DATE
Era	BIGINT
Longitude	DOUBLE
Latitude	DOUBLE
Enumeration	VARCHAR
Color	VARCHAR

Notice: derby does not support negative date, so Ear is saved as long.

3. All values in Matrix are saved as Double.

2.4.6 Calculate Data

1. Data are handled as original values, without concern about types and formats of columns.
2. Values are parsed as need. Example, if the calculation requires double values, then try to convert values as doubles.
3. If value conversion fails, then invalid value is handled as definition of column.
4. Calculation itself can define how to handle invalid values. Definition of calculation priors to definition of column.
5. Both column and calcultion can define decimal scale. Definition of calculation priors to definition of column.
6. Example, Column type is String, and descriptive statistic is running against it:
 - Try to convert each value as double.
 - If “count invalid value as zero”, then invalid values are calculated into “Count”, and participate in calculations of “Mean” and “Sum”.
 - If “skip invalid value”, then invalid values are not involved in any calculation.
 - If “count invalid value as empty”, then invalid values cause results of all calculations as invalid(Double.NaN).

2.4.7 Sort Data

1. For any calculations involved in sorting, data will be translated into a temporary database table, and be sorted by database system.
2. Results of sorting are related to columns' types. Example, string “124” is smaller than string “18”, while number “124” is bigger than number “18”.

2.4.8 Color of Column

Column color is mainly used for data charts.

When chart is generated, elements in it are displayed in colors as their columns' definitions.. Then user can set chart in random colors.

2.5 Edit Data

2.5.1 Mode "Table"

"Table" is the master edit mode:

1. Its modifications are applied to other panes automatically.
2. It is the final data to save.
3. To string values, except for text file, multiple lines can be edited and saved:
 - When the value is single line(not contain line break):
 - Text field is shown when click the data cell.
 - Write "\n" as line break in the value and commit the change(return or click other place).
 - When the value contains line breaks:
 - Text area is shown when click the data cell.
 - Write the text in multiple lines directly.

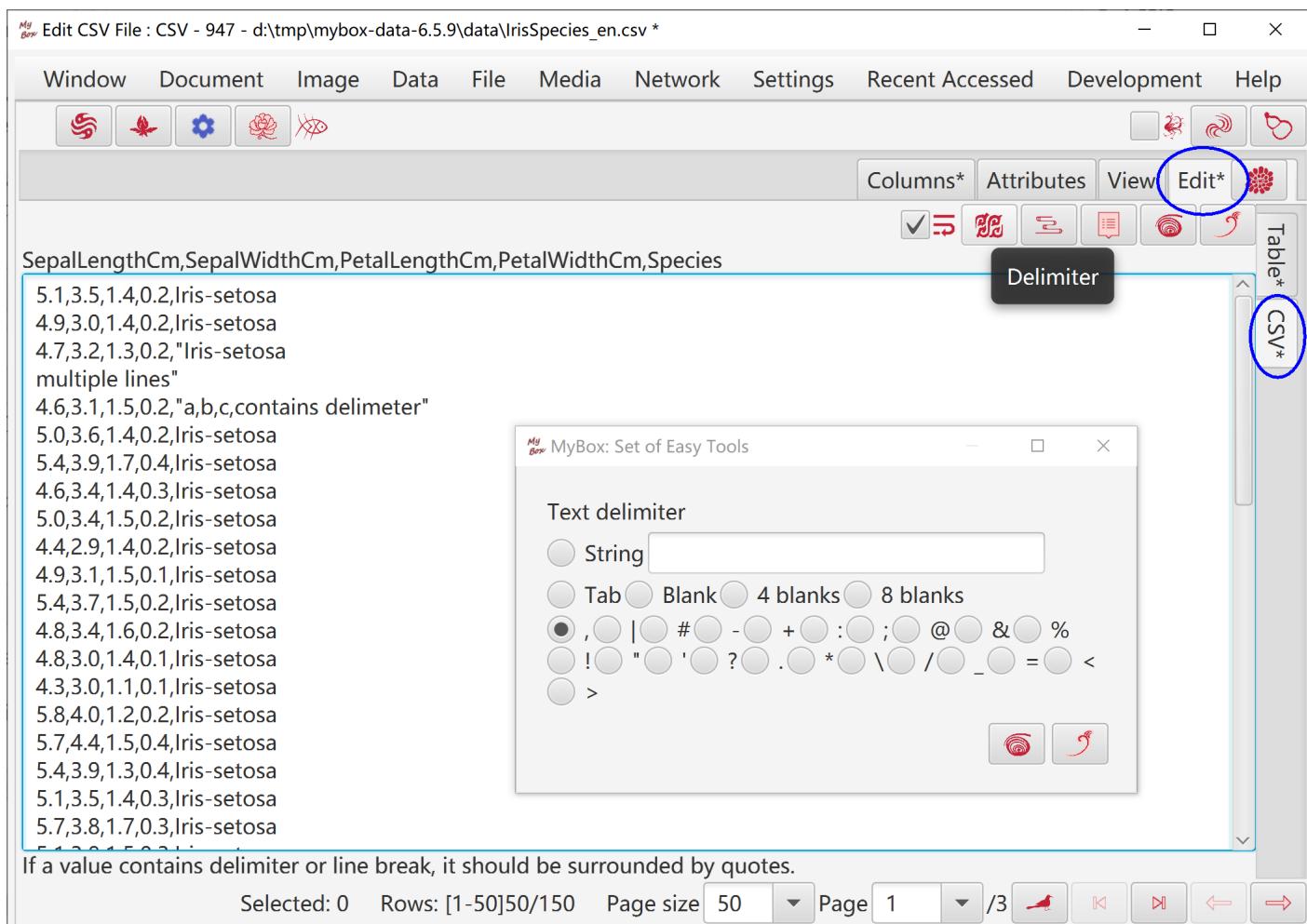
Table r...	Data...	SepalLeng...	SepalWidt...	PetalLeng...	PetalWidt...	Species
<input type="checkbox"/> 1	1	5.1	3.5	1.4	0.2	Iris-setosa
<input type="checkbox"/> 2	2	4.9	3	1.4	0.2	Iris-setosa
<input checked="" type="checkbox"/> 3	3	4.7	3.2	1.3	0.2	Iris-setosa multiple lines
<input type="checkbox"/> 4	4	4.6				Iris-setosa
<input type="checkbox"/> 5	5	5				
<input type="checkbox"/> 6	6	5.4				
<input type="checkbox"/> 7	7	4.6				
<input type="checkbox"/> 8	8	5				
<input type="checkbox"/> 9	9	4.4				
<input type="checkbox"/> 10	10	4.9				
<input type="checkbox"/> 11	11	5.4				
<input type="checkbox"/> 12	12	4.8				
<input type="checkbox"/> 13	13	4.8				
<input type="checkbox"/> 14	14	4.3				
<input type="checkbox"/> 15	15	5.8				

Selected: 0 Rows: [1-50]50/150 Page size 50 Page 1 /3

2.5.2 Mode "CSV"

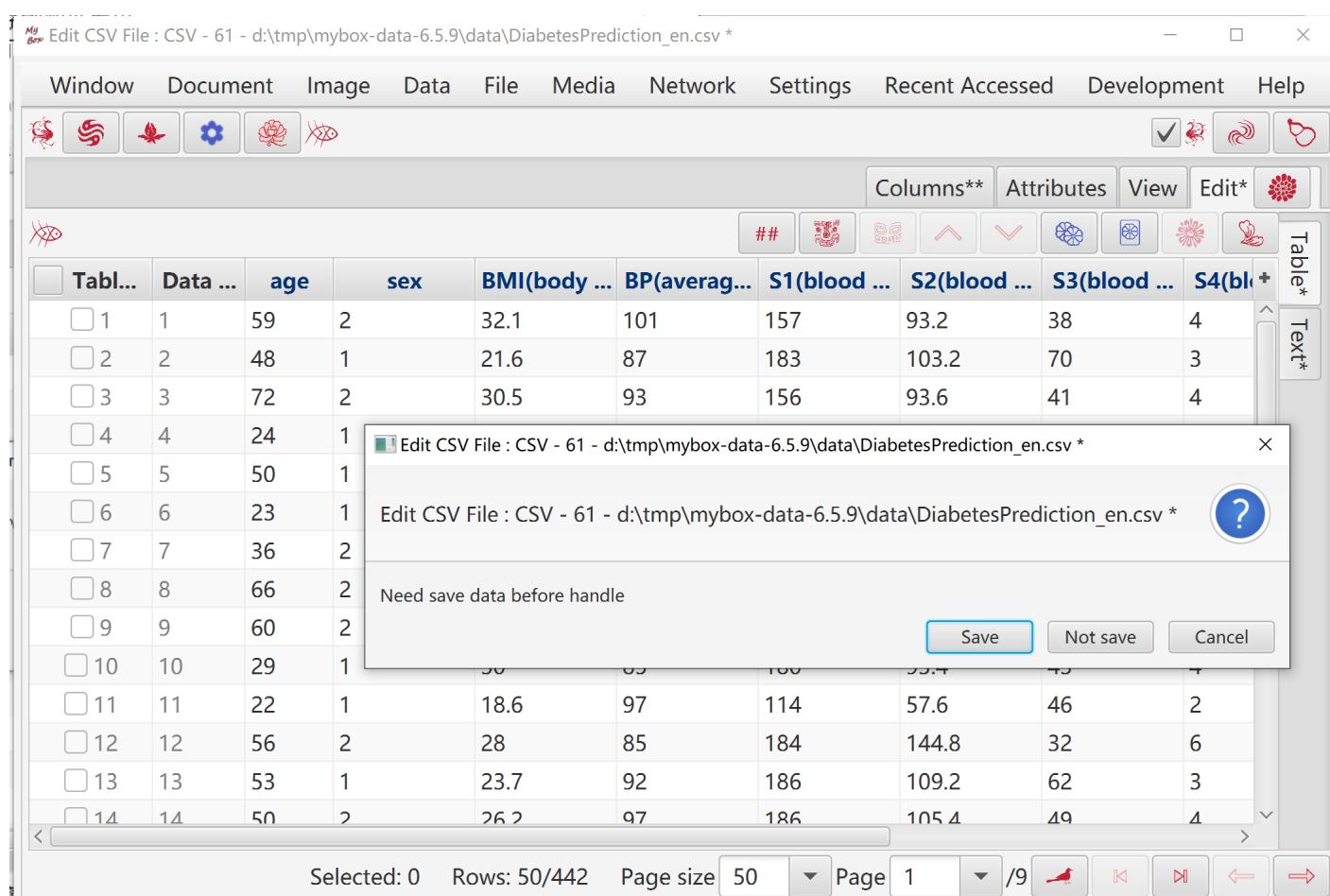
"CSV" is the assist edit mode.

1. Click button "OK" to apply its modifications to "Table".
2. Click button "Cancel" to discard its modifications and pick data from "Table".
3. Click button "Delimiter" to pick data from "Table" and apply new delimiter while its modifications are discarded.
4. If value contains delimiter or line break, it should be surrounded by quotes.



2.5.3 Save and Recover

1. When changes are not saved, * is displayed in tab header.
2. Click button "Save" to write modifications to file and database:
 - Changes of rows in "Table", including modify/add/delete/sort, affect rows of current page in file.
 - Changes in "Columns" tab, including modify/add/delete/sort, affect all rows in file.
 - Changes of attributes and columns are saved in database.
3. Click button "Recover" to discard all modifications and load data from file and database.



2.6 View Data

2.6.1 View Html

1. View data of current data.
2. Options: Form, Title, Column Names, Row Numbers.

MyBox Edit CSV File : CSV - 947 - d:\tmp\mybox-data-6.5.9\data\IrisSpecies_en.csv *

Window Document Image Data File Media Network Settings Recent Accessed Development Help

Form Title Column name Row number

Columns* Attributes **View** Edit*

d:\tmp\mybox-data-6.5.9\data\IrisSpecies_en.csv

Row number	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
Row1	5.1	3.5	1.4	0.2	Iris-setosa
Row2	4.9	3	1.4	0.2	Iris-setosa
Row3	4.7	3.2	1.3	0.2	Iris-setosa multiple lines
Row4	4.6	3.1	1.5	0.2	a,b,c,contains delimiter
Row5	5	3.6	1.4	0.2	Iris-setosa
Row6	5.4	3.9	1.7	0.4	Iris-setosa
Row7	4.6	3.4	1.4	0.3	Iris-setosa
Row8	5	3.4	1.5	0.2	Iris-setosa
Row9	4.4	2.9	1.4	0.2	Iris-setosa

H2

Selected: 0 Rows: [1-50]50/150 Page size 50 ▾ Page 1 ▾ /3

Html CSV

2.6.2 View CSV

1. View data of current data.
2. Options: Form, Title, Column Names, Row Numbers.
3. Set delimiter.

MyBox Edit CSV File : CSV - 947 - d:\tmp\mybox-data-6.5.9\data\IrisSpecies_en.csv *

Window Document Image Data File Media Network Settings Recent Accessed Development Help

Columns* Attributes View Edit* Form Title Column name Row number Delimiter

d:\tmp\mybox-data-6.5.9\data\IrisSpecies_en.csv

Delimiter

HTML CSV

multiple lines"

	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
1	5.1	3.5	1.4	0.2	Iris-setosa
2	4.9	3	1.4	0.2	Iris-setosa
3	4.7	3.2	1.3	0.2	"Iris-setosa
4	4.6	3.1	1.5	0.2	a,b,c,contains delimiter
5	5	3.6	1.4	0.2	Iris-setosa
6	5.4	3.9	1.7	0.4	Iris-setosa
7	4.6	3.4	1.4	0.3	Iris-setosa
8	5	3.4	1.5	0.2	Iris-setosa
9	4.4	2.9	1.4	0.2	Iris-setosa
10	4.9	3.1	1.5	0.1	Iris-setosa
11	5.4	3.7	1.5	0.2	Iris-setosa
12	4.8	3.4	1.6	0.2	Iris-setosa
13	4.8	3	1.4	0.1	Iris-setosa
14	4.3	3	1.1	0.1	Iris-setosa
15	5.8	4	1.2	0.2	Iris-setosa
16	5.7	4.4	1.5	0.4	Iris-setosa
17	5.4	3.9	1.3	0.4	Iris-setosa
18	5.1	3.5	1.4	0.3	Iris-setosa

Selected: 0 Rows: [1-50]50/150 Page size 50 Page 1 /3

MyBox: Set of Easy Tools

Text delimiter

String

Tab Blank 4 blanks 8 blanks

, | # - + : ; @ & %

! " ' ? . * \ / _ = <

>

Selected: 0 Rows: [1-50]50/150 Page size 50 Page 1 /3

2.7 Define Data Attributes

Data name, decimal scale, and maximum value of random.

The screenshot shows the MyBox Data Tools application window. The title bar reads "Edit CSV File : CSV - 7 - d:\tmp\mybox-data-6.5.9\data\SouthGermanCredit_en.csv". The menu bar includes "Window", "Document", "Image", "Data", "File", "Media", "Network", "Settings", "Recent Accessed", "Development", and "Help". The toolbar contains icons for various operations like Open, Save, Print, and Undo/Redo. The main area has tabs for "Columns" and "Attributes", with "Attributes" currently selected. Below the tabs, there are buttons for "View" and "Edit". A status bar at the bottom shows file information and navigation controls.

CSV - 7 - d:\tmp\mybox-data-6.5.9\data\SouthGermanCredit_en.csv

ID 7

Data type CSV

Data name South German credit

Decimal scale 2

Maximum value of random 1000

Description

http://archive.ics.uci.edu/ml/datasets/South+German+Credit
700 good and 300 bad credits with 20 predictor variables. Data from 1973 to 1975. Stratified sample from actual credits with bad credits heavily oversampled. A cost matrix can be used.

Update time 2022-09-28 11:17:37

File: d:\tmp\mybox-data-6.5.9\data\SouthGermanCredit_en.csv
File size: 46.934 KB
File modify time: 2022-09-28 11:17:35
Charset: UTF-8
Delimiter: ,
First line defines the columns' names: Yes
Lines number in file: 1000
Columns number: 21
Current page: 1 / 20
Rows range in page: 1 - 50 (50)
Page modify time: 2022-09-28 11:17:40

Selected: 0 Rows: 50/1000 Page size 50 Page 1 /20

2.8 Row Expression

2.8.1 Usages of Row Expression

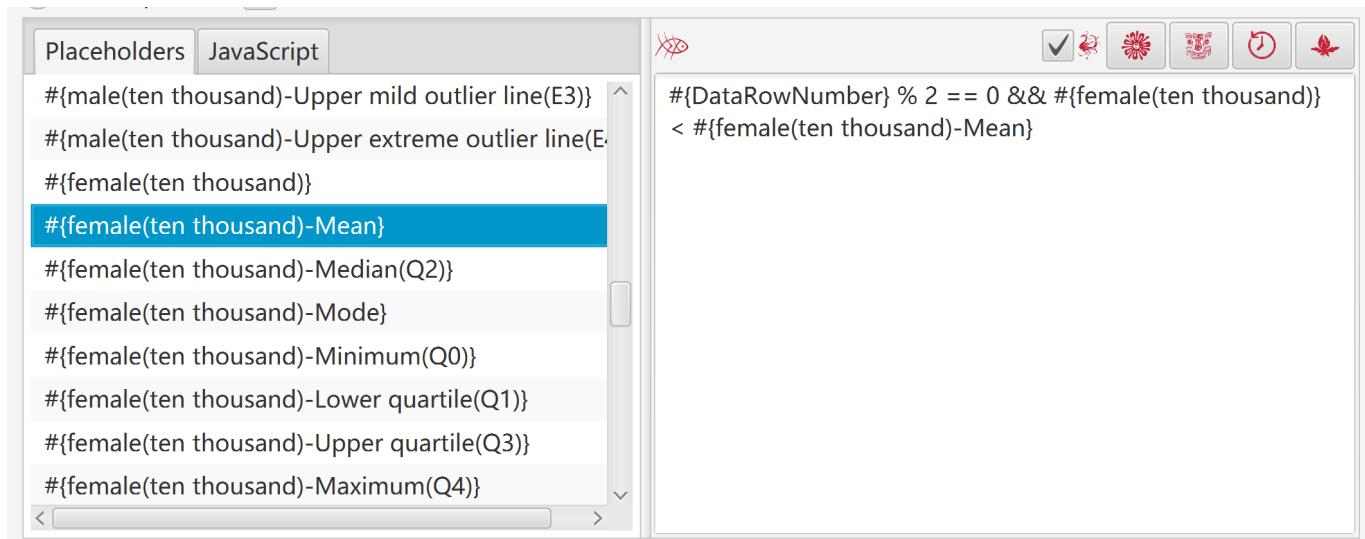
JavaScript expression can be data values when manufacture/trim/calculate data or generate chart.

2.8.2 Edit Row Expression

1. If the script is blank, then return empty string.
2. It can include any valid elements which Nashorn can parse(ECMAScript 5.1).
3. It should be a value finally.
4. It can include following placeholders:

```
#{{TableRowNumber}}
#{{DataRowNumber}}
#{{<column_name>}}
#{{<column_name>- <statistic-name>}}
```

5. All valid placeholders are listed in left.
6. To save the script, click "Edit" button.
7. Hover or click button “Examples” to paste example codes.
8. Hover or click button “Histories” to paste codes in histories.



2.8.3 Calculate Row Expression

When MyBox evaluates the expression:

1. Placeholders are replaced with actual values of each data row.
2. Statistic values are calculated by all data.
3. '#{xxx}' is handled as string while #{{xxx}} is handled as number.
4. When handles all pages, script fails when it includes “#{TableRowNumber}” .

2.8.4 Examples

expression	meaning
#DataRowNumber} % 2 == 0	data row number is even
#TableRowNumber} % 2 == 1	odd rows in current page
Math.abs(#{v1}) + Math.PI * Math.sqrt(#{v2})	calculation
'#{v1}'.replace(/hello/ig, 'Hello')	replace all “hello”(case-insensitive) as “Hello” in column “v1”
'#{v1}'.toLowerCase()	lower case of value of column “v1”
'#{v1}'.split(',')	split value of column “v1” by comma
#{v1} - #{v1-Mean}	difference between value of column “v1” and mean of column “v1”
new Date('#{time}'.replace(/-/g,'/')).getFullYear()	year of value of column “time”

2.9 Row Filter

2.9.1 Usages of Row Filter

“Row Filter” is special “Row Expression”, and can be condition to filter data rows.

2.9.2 Edit Row Filter

1. It should be boolean value("true" or "false") finally.
2. Can set maximum rows to take.
3. Can be saved in tree.

Maximum filtered rows to take(Empty/zero/negative to unlimit)

Select to paste

Placeholders Saved

Take source data when the expression is blank or
 true Others

'#{{Column1}} == ''

Value is null or empty

'#{{Column1}}.length > 0'

2.9.3 Examples

Expression	meaning
<code>#{{DataRowNumber}} % 2 == 0</code>	data row number is even
<code>#{{TableRowNumber}} % 2 == 1</code>	odd rows in current page
<code>Math.abs(#{{v1}}) >= 0</code>	value of column “v1” is number
<code>#{{v1}} > 0</code>	value of column “v1” is larger than zero
<code>#{{v1}} - #{{v2}} < 100</code>	difference between values of “v1” and “v2” is less than 100
<code>'#{{v1}}' == ''</code>	value of column “v1” is null or empty
<code>'#{{v1}}'.length > 0</code>	value of column “v1” is not empty
<code>'#{{v1}}'.search(/Hello/ig) >= 0</code>	value of column “v1” includes “Hello”(case-insensitive)
<code>'#{{v1}}'.startsWith('Hello')</code>	value of column “v1” starts with “Hello”
<code>var array = ['A', 'B', 'C'];array.includes('#{{v1}}')</code>	value of column “v1” is one of “A”, “B”, “C”
<code>#{{v1}} < #{{v1-Mean}}</code>	value of column “v1” is less than mean of column “v1”
<code>new Date('#{{time}}'.replace(/-/g,'/')).getTime() > new Date('2016/05/19 09:23:12').getTime()</code>	value of column “time” is later than '2016/05/19 09:23:12'

2.10 Group Rows

2.10.1 Before and after Grouping

1. Before grouping, rows can be filtered. That is, objects of grouping are filtered rows.
2. In order to implement sorting and statistic with database system, data are converted as temporary database table automatically before grouping.
3. Sorting condition and maximum data number of each group can be set for grouped data.

2.10.2 Usages of grouping

With grouped data, following can be handled:

1. Add grouping information in source data, which are 2 new columns: Group Number and Group Parameter.
2. Split data into multiple files of groups.
3. Calculate descriptive statistic of groups.
4. Display following types of charts in sequence of groups dynamically: XY chart, pie chart, comparison bars chart, self comparison bars chart, box-and-whisker chart.

2.10.3 Ways of Data Grouping

MyBox supports six ways to group data:

Grouping		Results	
Way	Parameters	complete	Non-intersect
Equal values	Several column names	Yes	Yes
Value Ranges	Column name + Size of split	Yes	Yes
	Column name + Number of split	Yes	Yes
	Column name + List of “start-end”	Possible	Possible
Time values	Column of date/time + Time unit	Yes	Yes
Expression	Row expression	Yes	Yes
Conditions	Several row filters	Possible	Possible
Row numbers	Size of split	Yes	Yes
	Number of split	Yes	Yes
	List of “Start-end”	Possible	Possible

Following are their descriptions.

2.10.4 Group by Equal Values

2.10.4.1 Way of Grouping

Select several columns and put rows into same group when they have same values of these columns.

The screenshot shows the 'Split/Group' dialog box with the 'Group' tab selected. The interface includes a toolbar with buttons for Data, Row filter, Group, and Options. Below the toolbar, there are seven radio buttons labeled 'Equal values', 'Value range', 'Time', 'Expression', 'Conditions', and 'Rows range'. Underneath these buttons are three icons: a double up arrow, a double down arrow, and a double up-down arrow, followed by the text 'Group by'. A table below lists columns from 1 to 11. Columns 1 ('age') and 2 ('sex') have checkboxes next to them, which are checked. Other columns (3 through 11) have empty checkboxes. The table has columns for 'Column' and row numbers.

	Column	
1	age	
2	sex	
3	BMI(body mass index)	
4	BP(average blood pressure)	
5	S1(blood serum measurement 1)	
6	S2(blood serum measurement 2)	
7	S3(blood serum measurement 3)	
8	S4(blood serum measurement 4)	
9	S5(blood serum measurement 5)	
10	S6(blood serum measurement 6)	
11	disease progression one year after baseline	

Groups are generated against the filtered data rows as following:

- Select some columns.
- Rows are divided into same group when they have equal values of these columns.

2.10.4.2 Implementation of Grouping

1. Convert data as a temporary database table
2. Read each row and group them by equal values of selected columns. The results like executing SQL of “group by”.

2.10.4.3 Results of Grouping

1. Complete: Each row belongs to a group.
 2. Non-intersect: Each row only belongs to one group.

My Box Edit CSV File : CSV - 949 - d:\tmp\mybox-data-6.5.9\generated\Progression_Prediction_of_Diabetes_Equal_values_age_sex.csv

Table r...	Data...	Group	Equal values_age_sex	age	sex	BMI(body ...	BP(averag...
<input type="checkbox"/> 1	1	1	{sex=1, age=19.0}	19	1	23.2	75
<input type="checkbox"/> 2	2	1	{sex=1, age=19.0}	19	1	25.3	83
<input type="checkbox"/> 3	3	1	{sex=1, age=19.0}	19	1	19.2	87
<input type="checkbox"/> 4	4	2	{sex=1, age=20.0}	20	1	22.9	87
<input type="checkbox"/> 5	5	3	{sex=2, age=20.0}	20	2	24.2	88
<input type="checkbox"/> 6	6	3	{sex=2, age=20.0}	20	2	22.1	87
<input type="checkbox"/> 7	7	4	{sex=1, age=21.0}	21	1	24.2	76
<input type="checkbox"/> 8	8	4	{sex=1, age=21.0}	21	1	20.1	63
<input type="checkbox"/> 9	9	5	{sex=1, age=22.0}	22	1	19.9	75
<input type="checkbox"/> 10	10	5	{sex=1, age=22.0}	22	1	19.3	82
<input type="checkbox"/> 11	11	5	{sex=1, age=22.0}	22	1	23	73
<input type="checkbox"/> 12	12	5	{sex=1, age=22.0}	22	1	18.6	97
<input type="checkbox"/> 13	13	6	{sex=1, age=23.0}	23	1	18.8	78
<input type="checkbox"/> 14	14	6	{sex=1, age=23.0}	23	1	29	90
<input type="checkbox"/> 15	15	6	{sex=1, age=23.0}	23	1	22.6	89
<	16	7	{sex=1, age=23.0}	23	2	18	78
>							

Selected: 0 Rows: [1-50]50/442 Page size 50 Page 1 /9

2.10.5 Group by Value Ranges

2.10.5.1 Way of Grouping

Select one column and divide data by value ranges of this column.

Following 3 selections can be made for the split:

- Interval of values
- Number of split
- List of "StartValue-EndValue"

The screenshot shows the 'Split/Group' dialog box for a CSV file named 'DiabetesPrediction_2023-01-19...'. The top bar has standard window controls. The main title is 'Split/Group'. Below it is a toolbar with tabs: 'Data' (selected), 'Row filter', 'Group', and 'Options'. The 'By' section is set to 'Value range'. The 'Column' dropdown is set to 'age'. Under 'Value range' settings, 'Number of split' is selected with the value '3'. The 'Decimal scale of ranges' is set to '2'. At the bottom right are two small icons: a red swirl and a red musical note.

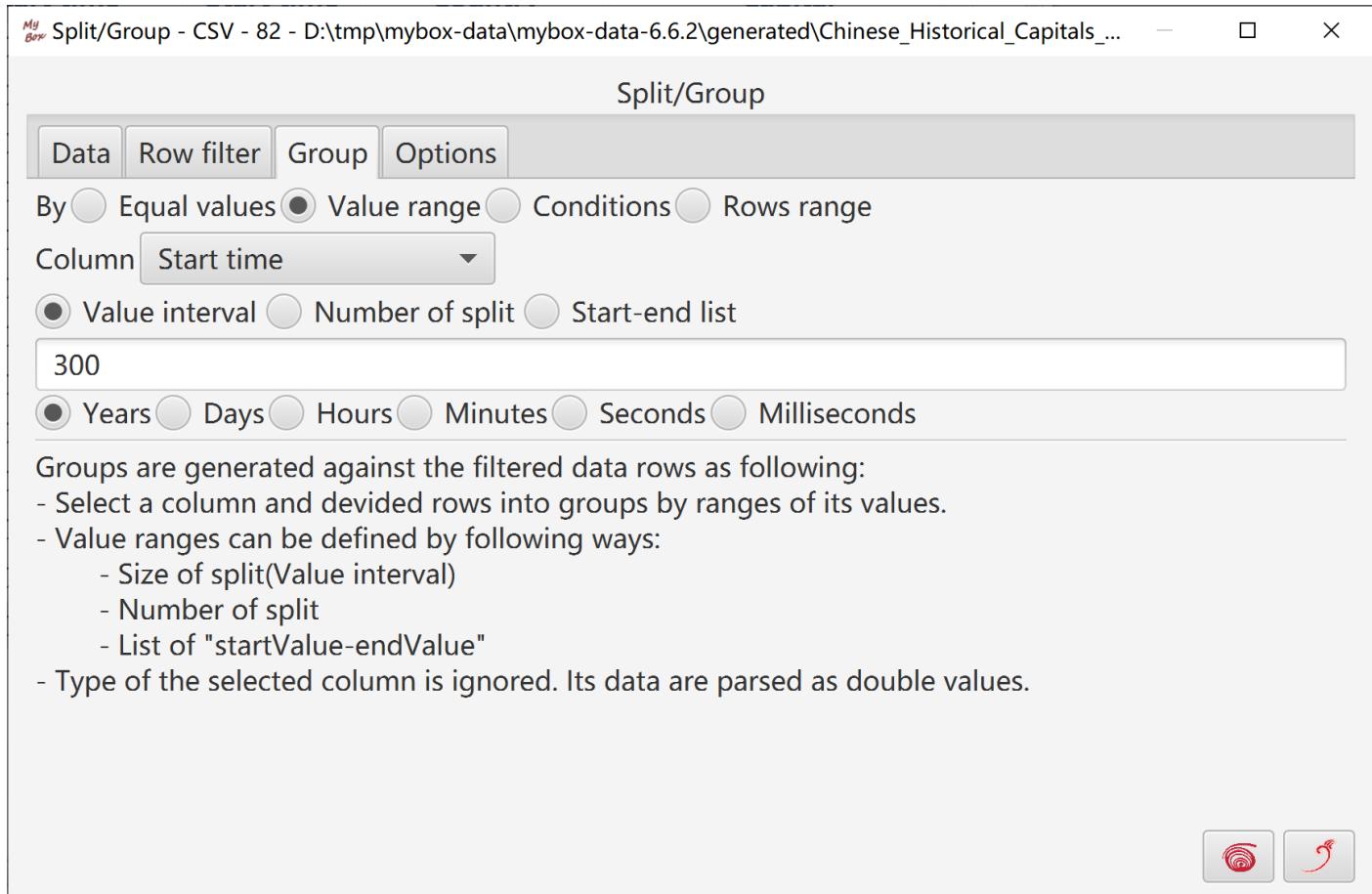
Groups are generated against the filtered data rows as following:

- Select a column and divide rows into groups by ranges of its values.
- Value ranges can be defined by following ways:
 - Size of split(Value interval)
 - Number of split
 - List of "startValue-endValue"
- Type of the selected column is ignored. Its data are parsed as double values.

2.10.5.2 Split Date/Era

When split rows of date/era:

1. Values of the column are converted as milliseconds.
2. Time unit can be set.
3. For items of “Start-End”, ranges should be in valid formats of date/era.



2.10.5.3 Edit List of “Start-End”

1. Add/Delete/Update start-end items in table.
2. Set whether includes “start”/“end” for “start-end” items.

MyBox Split/Group - CSV - 82 - D:\tmp\mybox-data\mybox-data-6.6.2\generated\Chinese_Historical_Capitals_...

Split/Group

Data Row filter Group Options

By Equal values Value range Conditions Rows range

Column **Start time**

Value interval Number of split Start-end list

Start	End	Include "start"	Include "end"	
2022-02-01 00:00:00	2022-08-30 00:00:00	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
2022-05-30 16:23:08	2022-11-30 16:23:12		<input checked="" type="checkbox"/>	
2022-12-01 00:00:00	2022-12-31 00:00:00	<input checked="" type="checkbox"/>		

Groups are generated against the selected column.

- Select a column and devide rows into groups.
- Value ranges can be defined by
 - Size of split(Value interval)
 - Number of split
 - List of "startValue-endValue"
- Type of the selected column is

Value range

Column: Start time

Start  Include ...

End  Include ...

2.10.5.4 Implementation of Grouping

- Convert data as a temporary database table.
- Ignore type of selected column, and parse its values as double.
- Calculate maximum value and minimum value of this column.
- Count ranges of each group.
- Compose query statements for each group. The conditions are like “column-name \geq start-value-of-range AND column-name $<$ end-value-of-range”.

2.10.5.5 Results of Grouping

- If split as size or number, then:
 - Complete: Each row belongs to a group.
 - Non-intersect: Each row only belongs to one group.
- If split as start-end list, then:
 - Possible incomplete: Some rows may not belong to any group.
 - Possible intersect: Some rows may belong to several groups.

MyBox Edit CSV File : CSV - 951 - d:\tmp\mybox-data-6.5.9\generated\Progression_Prediction_of_Diabetes_Range_age.csv

Window Document Image Data File Media Network Settings Recent Accessed Development Help

Columns* Attributes View Edit

Table CSV

Selected: 0 Rows: [1-50]50/442 Page size 50 Page 1 /9

Table r...	Data...	Group	Range_age	age	sex	BMI(body ...	BP(averag...	S1(blood ...	S2(blood ..
<input type="checkbox"/>	20	1	[19,27)	20	2	24.2	88	126	72.2
<input type="checkbox"/>	21	1	[19,27)	26	1	18.8	83	191	103.6
<input type="checkbox"/>	22	1	[19,27)	23	2	18	78	171	96
<input type="checkbox"/>	23	1	[19,27)	19	1	25.3	83	225	156.6
<input type="checkbox"/>	24	1	[19,27)	25	2	22.6	85	130	71
<input type="checkbox"/>	25	1	[19,27)	19	1	23.2	75	143	70.4
<input type="checkbox"/>	26	1	[19,27)	25	2	23.5	88	143	80.8
<input type="checkbox"/>	27	2	[27,35)	29	1	30	85	180	93.4
<input type="checkbox"/>	28	2	[27,35)	34	2	24.7	118	254	184.2
<input type="checkbox"/>	29	2	[27,35)	31	1	29.7	88	167	103.4
<input type="checkbox"/>	30	2	[27,35)	30	2	25.2	83	178	118.4
<input type="checkbox"/>	31	2	[27,35)	32	1	30.5	89	182	110.6
<input type="checkbox"/>	32	2	[27,35)	32	2	25.4	90.33	153	100.4
<input type="checkbox"/>	33	2	[27,35)	33	1	25.3	85	155	85
<input type="checkbox"/>	34	2	[27,35)	27	1	19.6	78	128	68

2.10.6 Group by Time

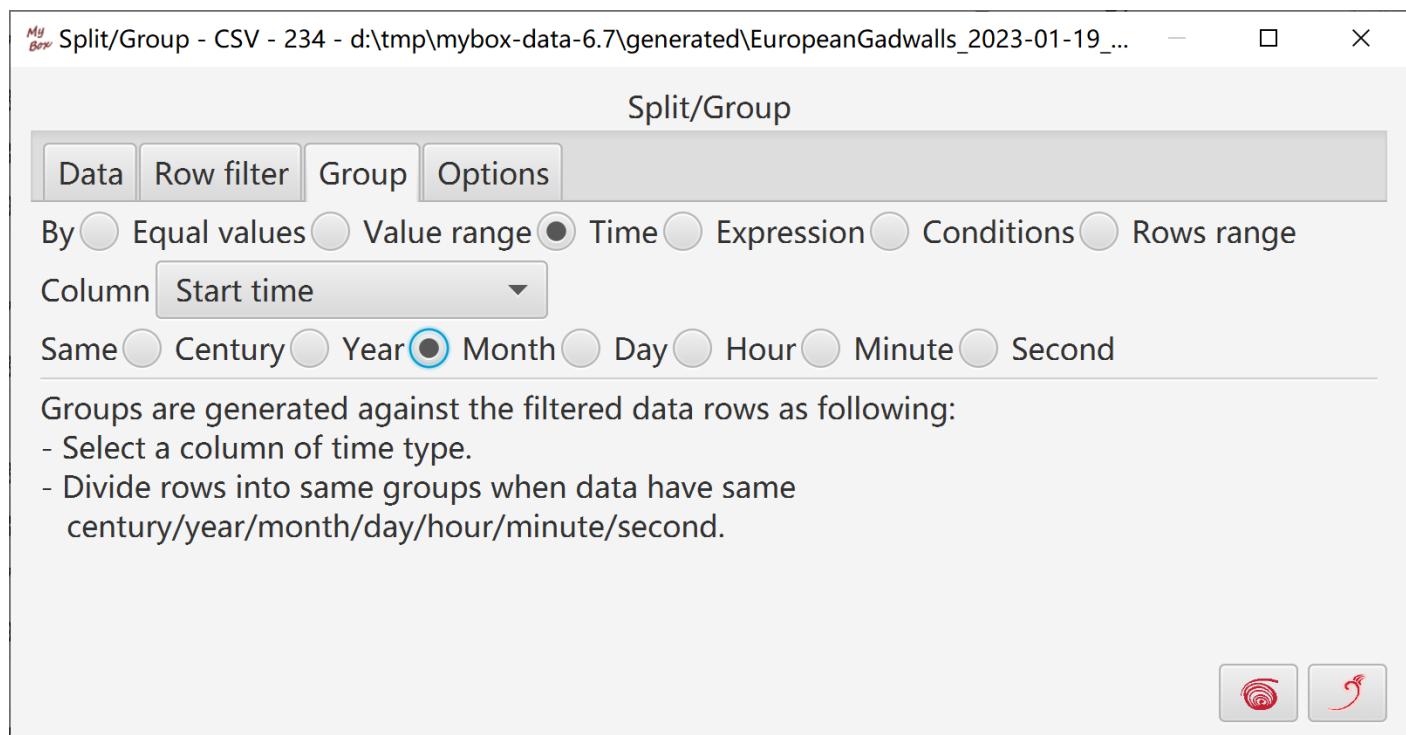
2.10.6.1 Way of Grouping

Only when data include date/time columns, this way can be taken:

1. Select a column of date/time type.
2. Select time unit: century/year/month/day/hour/minute/second.

Examples:

- When time unit is “century”, data are calculated to get century values and grouped by them.
- When time unit is “month”, data are divided as same groups when they have same values of year and month.
- When time unit is “hour”, data are divided as same groups when they have same values of year, month, day, and hour.
- When time unit is “second”, data are divided as same groups when they have same values of year, month, day, hour, minute, and second.



2.10.6.2 Implementation of Grouping

1. Convert data as a temporary database table.
2. As time unit, save values are in same groups.

2.10.6.3 Results of Grouping

1. Complete: Each row belongs to a group.
2. Non-intersect: Each row only belongs to one group.

Edit CSV File : CSV - 240 - d:\tmp\mybox-data-6.7\generated\Autumn_movement_patterns_of_European_Gadwalls_Month_Start_time.csv

Window Document Image Data File Media Network Settings Recent Accessed Development Help

Columns* Attributes View Edit Table CSV

Table CSV

Selected: 0 Rows: [2451-2500] 50/2508 Page size 50 Page 50 /51

Table row	Data row	Group id	Month_Start time	Start time	End time	Longitude	Lat
15	2465	16	2010-10	2010-10-02 15:00:00	2010-10-02 15:00:00	12.678	48.70
16	2466	16	2010-10	2010-10-02 09:00:00	2010-10-02 09:00:00	12.67933	48.70
17	2467	16	2010-10	2010-10-02 06:00:00	2010-10-02 06:00:00	12.67783	48.70
18	2468	16	2010-10	2010-10-01 21:00:00	2010-10-01 21:00:00	12.6775	48.70
19	2469	16	2010-10	2010-10-01 15:00:00	2010-10-01 15:00:00	12.67933	48.70
20	2470	16	2010-10	2010-10-01 09:00:00	2010-10-01 09:00:00	12.67783	48.70
21	2471	16	2010-10	2010-10-01 06:00:00	2010-10-01 06:00:00	12.67783	48.70
22	2472	17	2010-11	2010-11-27 12:38:00	2010-11-27 12:38:00	11.123	48.74
23	2473	17	2010-11	2010-11-18 17:16:00	2010-11-18 17:16:00	11.041	48.72
24	2474	17	2010-11	2010-11-18 15:00:00	2010-11-18 15:00:00	11.08567	48.73
25	2475	17	2010-11	2010-11-18 09:00:00	2010-11-18 09:00:00	11.0855	48.73
26	2476	17	2010-11	2010-11-17 21:00:00	2010-11-17 21:00:00	11.06883	48.73
27	2477	17	2010-11	2010-11-17 15:00:00	2010-11-17 15:00:00	11.0855	48.73
28	2478	17	2010-11	2010-11-17 09:00:00	2010-11-17 09:00:00	11.08102	48.73

2.10.7 Group by Expression

2.10.7.1 Way of Grouping

Define a row expression, and rows are in same groups when they have same values of this expression.

Examples:

- Expression “new Date("#{StartTime}).getDay()”, group as the week values of column “StartTime”.
- Expression “#{description}.charAt(2)”, group as the 2th character of column “description”.
- Expression “#{StartTime}.substring(0, 7)”, group as the first 7 characters of column “StartTime”.

The screenshot shows the 'Split/Group' dialog box with the 'Group' tab selected. The 'Expression' radio button is selected. The expression '#{Start time}.substring(0, 7)' is entered in the main area. A preview pane shows icons for eye, checkmark, flower, bell, and cloud.

Groups are generated against the filtered data rows as following:

- Divide rows into same groups when their results of the row expression are same.

2.10.7.2 Implementation of Grouping

1. Convert data as a temporary database table.
2. Calculate the expression against each row, save values are in same groups.

2.10.7.3 Results of Grouping

1. Complete: Each row belongs to a group.
2. Non-intersect: Each row only belongs to one group.

MyBox Edit CSV File : CSV - 243 - d:\tmp\mybox-data-6.7\generated\Autumn_movement_patterns_of_European_Gadwalls_Expression.csv

Table row	Data row	Group id	Expression	Start time	End time	Longitude	Latitude
16	2466	16	2010-10	2010-10-02 09:00:00	2010-10-02 09:00:00	12.67933	48.70383
17	2467	16	2010-10	2010-10-02 06:00:00	2010-10-02 06:00:00	12.67783	48.7035
18	2468	16	2010-10	2010-10-01 21:00:00	2010-10-01 21:00:00	12.6775	48.7046
19	2469	16	2010-10	2010-10-01 15:00:00	2010-10-01 15:00:00	12.67933	48.70383
20	2470	16	2010-10	2010-10-01 09:00:00	2010-10-01 09:00:00	12.67783	48.7035
21	2471	16	2010-10	2010-10-01 06:00:00	2010-10-01 06:00:00	12.67783	48.7035
22	2472	17	2010-11	2010-11-27 12:38:00	2010-11-27 12:38:00	11.123	48.744
23	2473	17	2010-11	2010-11-18 17:16:00	2010-11-18 17:16:00	11.041	48.729
24	2474	17	2010-11	2010-11-18 15:00:00	2010-11-18 15:00:00	11.08567	48.73083
25	2475	17	2010-11	2010-11-18 09:00:00	2010-11-18 09:00:00	11.0855	48.73083
26	2476	17	2010-11	2010-11-17 21:00:00	2010-11-17 21:00:00	11.06883	48.73333
27	2477	17	2010-11	2010-11-17 15:00:00	2010-11-17 15:00:00	11.0855	48.73083
28	2478	17	2010-11	2010-11-17 09:00:00	2010-11-17 09:00:00	11.08483	48.7305

Selected: 0 Rows: [2451-2500]50/2508 Page size 50 /51

2.10.8 Group by Conditions

2.10.8.1 Way of Grouping

Define some row filters and data rows are divided in groups by these row filters.

The screenshot shows the 'Split/Group' dialog box from the MyBox Data Tools. The title bar reads 'Split/Group - CSV - 942 - d:\tmp\mybox-data-6.5.9\data\DiabetesPrediction_en.csv'. The main area is titled 'Split/Group' and contains a toolbar with four buttons: 'Data', 'Row filter' (which is selected), 'Group', and 'Options'. Below the toolbar, there's a section labeled 'By' with three radio button options: 'Equal values', 'Value range', and 'Conditions' (which is selected). There are also three icons: a red double arrow, a red asterisk, and a red rose. A table below lists six row filters:

Expression	Reverse	Maxim...	
<code>#{age} < 20</code>	false	-1	
<code>#{age} < 45</code>	false	-1	
<code>#{age} < 60</code>	false	-1	
<code>#{age} < 70</code>	false	-1	
<code>#{age} < 85</code>	false	-1	
<code>#{age} > 85</code>	false	-1	

At the bottom right of the dialog are two icons: a red swirl and a red heart.

Groups are generated against the filtered data rows as following:

- Define some row filters.
- Rows are divided in groups by these row filters.

2.10.8.2 Implementation of Grouping

3. Convert data as a temporary database table.
4. For each row filter, scan data rows to pick out group members which satisfy condition.

2.10.8.3 Results of Grouping

- Possible incomplete: Some rows may not belong to any group.
- Possible intersect: Some rows may belongs to several groups.

MyBox Edit CSV File : CSV - 953 - d:\tmp\mybox-data-6.5.9\generated\Progression_Prediction_of_Diabetes_Condition.csv

Table r...	Data...	Group	Condition	age	sex	BMI(body ...	BP(averag...
1	1	1	<code>#{age} < 20</code>	19	1	19.2	87
2	2	1	<code>#{age} < 20</code>	19	1	25.3	83
3	3	1	<code>#{age} < 20</code>	19	1	23.2	75
4	4	2	<code>#{age} < 45</code>	24	1	25.3	84
5	5	2	<code>#{age} < 45</code>	23	1	22.6	89
6	6	2	<code>#{age} < 45</code>	36	2	22	90
7	7	2	<code>#{age} < 45</code>	29	1	30	85
8	8	2	<code>#{age} < 45</code>	22	1	18.6	97
9	9	2	<code>#{age} < 45</code>	34	2	24.7	118
10	10	2	<code>#{age} < 45</code>	38	1	25.4	84
11	11	2	<code>#{age} < 45</code>	41	1	24.7	83
12	12	2	<code>#{age} < 45</code>	35	1	21.1	82
13	13	2	<code>#{age} < 45</code>	25	2	24.3	95
14	14	2	<code>#{age} < 45</code>	25	1	26	92
15	15	2	<code>#{age} < 45</code>	31	1	29.7	88
16	16	2	<code>#{age} < 45</code>	20	2	25.0	82

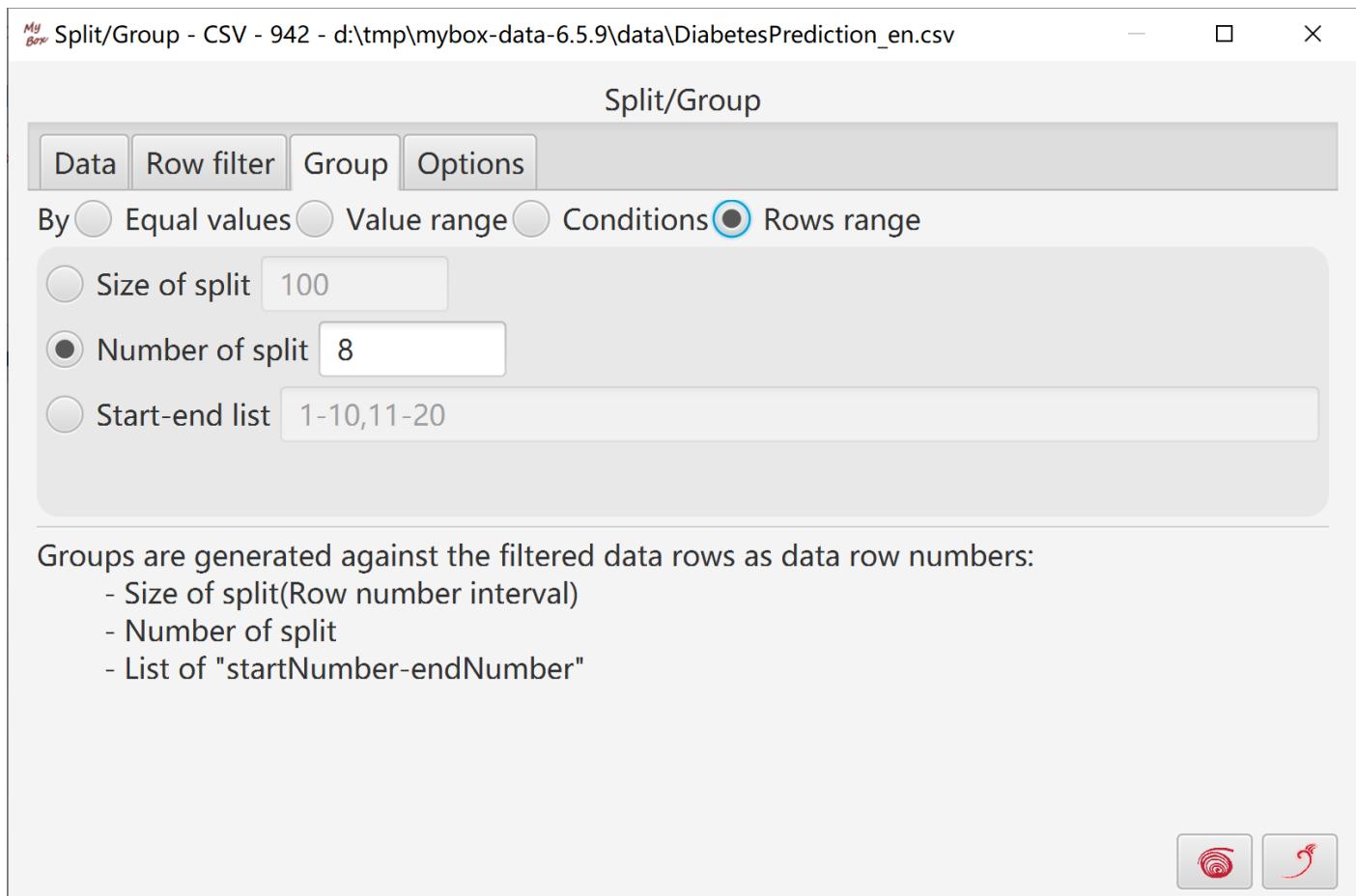
Selected: 0 Rows: [1-50]50/1374 Page size 50 Page 1 /28

2.10.9 Group by Row Number

2.10.9.1 Way of Grouping

Divide data rows by their row numbers. Following 3 selections can be for the split:

- Size of split(interval of row numbers)
- Number of split
- List of "StartNumber-EndNumber"



2.10.9.2 Implementation of Grouping

1. Convert data as a temporary database table.
2. Count range of row numbers for each group.
3. Scan data rows, and set groups as row numbers.

2.10.9.3 Results of Grouping

- If split as size or number, then:
 - Complete: Each row belongs to a group.
 - Non-intersect: Each row only belongs to one group.
- If split as start-end list, then:
 - Possible incomplete: Some rows may not belong to any group.
 - Possible intersect: Some rows may belong to several groups.

MyBox Edit CSV File : CSV - 955 - d:\tmp\mybox-data-6.5.9\generated\Progression_Prediction_of_Diabetes_DataRowNumber.csv

Table r...	Data...	Group	DataRowNumber	age	sex	BMI(body ...	BP(averag...
<input type="checkbox"/>	1	1	[1,55]	58	1	25.7	99
<input type="checkbox"/>	2	1	[1,55]	65	2	27.9	103
<input type="checkbox"/>	3	1	[1,55]	34	1	25.5	93
<input type="checkbox"/>	4	1	[1,55]	46	1	24.9	115
<input type="checkbox"/>	5	1	[1,55]	35	1	28.7	97
<input type="checkbox"/>	6	2	[56,110]	37	1	21.8	84
<input type="checkbox"/>	7	2	[56,110]	37	1	30.2	87
<input type="checkbox"/>	8	2	[56,110]	41	1	20.5	80
<input type="checkbox"/>	9	2	[56,110]	60	1	20.4	105
<input type="checkbox"/>	10	2	[56,110]	66	2	24	98
<input type="checkbox"/>	11	2	[56,110]	29	1	26	83
<input type="checkbox"/>	12	2	[56,110]	37	2	26.8	79
<input type="checkbox"/>	13	2	[56,110]	41	2	25.7	83
<input type="checkbox"/>	14	2	[56,110]	39	1	22.9	77
<input type="checkbox"/>	15	2	[56,110]	67	2	24	83

Selected: 0 Rows: [51-100]50/442 Page size 50 Page 2 /9

2.11 Modify Data

Hover or click button “Function” to select functions under menu item “Modify”.

The screenshot shows the MyBox Data Tools interface for editing a CSV file named "DiabetesPrediction_en.csv". The window title bar reads "Edit CSV File : CSV - 942 - d:\tmp\mybox-data-6.5.9\data\DiabetesPrediction_en.csv". The menu bar includes Window, Document, Image, Data, File, Media, Network, Settings, Recent Accessed, Development, and Help. The toolbar features icons for various functions like Add rows, Set values, Delete, Set Styles, Paste content, Trim, Calculation, Charts, Group Charts, and Examples. A context menu is open over a table, with the "Modify" option highlighted. The table contains 15 rows of data from the Diabetes dataset. The bottom of the interface shows navigation controls for selected rows (0-50), page size (50), and page number (1/9).

Index	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BreastDensity	Age
1	1	59					90
2	2	48					11
3	3	72					83
4	4	24					85
5	5	50					180
6	6	23					97
7	7	36	2	22	90.5	119.4	42
8	8	66	2	26.2	11	184	43
9	9	60	2	32.1	83	114	57.6
10	10	29	1	30	85	186	46
11	11	22	1	18.6	97	144.8	32
12	12	56	2	28	85	186	109.2
13	13	53	1	23.7	92	105.4	62
14	14	50	2	26.2	97	115.4	49
15	15	61	1	24	91	202	72

Selected: 0 Rows: [1-50]50/442 Page size 50 Page 1 /9

2.11.1 Add Rows

1. Select where to add new rows: Front, End, Above some row, Below some row.
2. Set rows number to be added.
3. Set values for new rows. Edit controls are provided as types of columns:
 - o If data contains both Longitude and Latitude, display button “Location”.
 - o For datetime, date, or era, display button “Example”.
 - o For boolean, display selection buttons.
 - o For enumeration, display value list.
 - o For color, display button “Color”.

Edit CSV File

Add rows

Front End Above Below

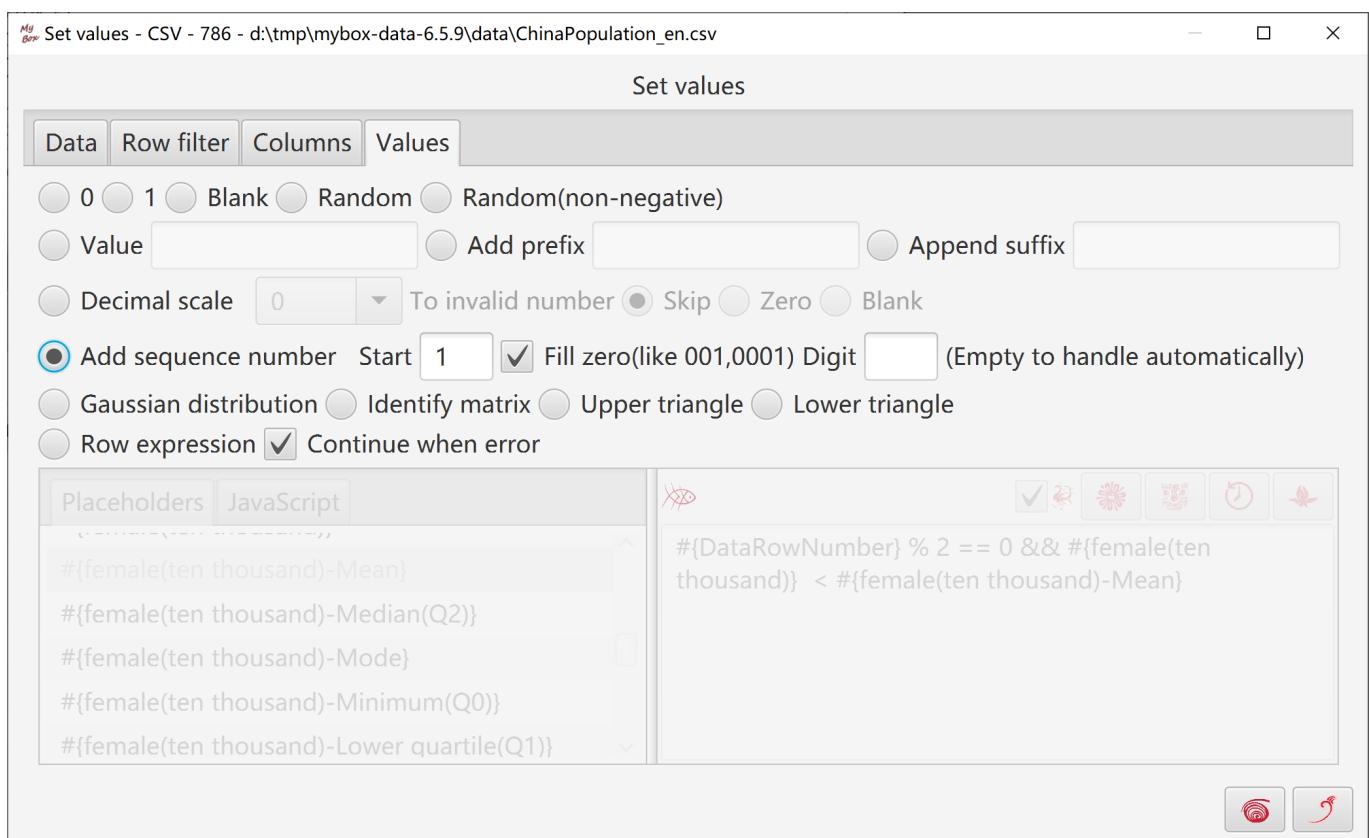
Table row

Rows number

DataRowNumber	<input type="button" value="Location"/>
Start time	2022-11-14
End time	2022-11-14
color	<input type="color" value="#FF0000"/> <input type="button" value="Color"/>
marked	<input type="radio"/> true <input checked="" type="radio"/> false
Longitude	115.3125
Latitude	30.14513
Altitude	
Coordinate System	CGCS2000(China Geodetic Coordinate System) <input type="button" value="Example"/>
Comments	CGCS2000(China Geodetic Coordinate System)
	GCJ-02(China encrypted coordinate)
	WGS-84(GPS)
	BD-09(Baidu encrypted coordinate)
	Mapbar coordinate

2.11.2 Set Values

1. Select data rows. Row filter can be set.
2. Select columns to be set.
3. Select way to set values:
 - o Constant: 0, 1, blank, or inputted value
 - o Random, random of non-negative
 - o Add prefix, append suffix, set decimal scale, or add sequence numbers
 - o When selected data are square, whose rows number equals to columns number, they can be set as following: gaussian distribution, identify matrix, upper triangular matrix, lower triangular matrix.
 - o Row expression
4. If handle all pages of data file, then auto-backup before set values.



2.11.3 Delete

1. Select data rows. Row filter can be set.
2. Option: Whether continue when error.
3. If handle all pages of data file, then auto-backup before delete.

MyBox Delete - CSV - 786 - d:\tmp\mybox-data-6.5.9\data\ChinaPopulation_en.csv

Delete

Data Row filter Options

Rows Selected Current page All pages 

<input type="checkbox"/> Table row	<input checked="" type="checkbox"/> Data row	Select some to handle, or select none to handle all in table: n(te... rural(ten ...						
<input type="checkbox"/> 1	<input checked="" type="checkbox"/>	1	167	28,145	26,022	5,765	48,402	
<input type="checkbox"/> 2	<input checked="" type="checkbox"/>	2	1950	55,106	38,669	26,527	6,169	1,027
<input type="checkbox"/> 3	<input checked="" type="checkbox"/>	3	1951	52,896	30,168	28,328	6,169	1,027
<input type="checkbox"/> 4	<input checked="" type="checkbox"/>	4	1952	57,482	29,833	27,649	7,163	50,319
<input type="checkbox"/> 5	<input checked="" type="checkbox"/>	5	1953	52,896	30,168	28,328	6,169	1,027
<input type="checkbox"/> 6	<input checked="" type="checkbox"/>	6	1954	57,482	29,833	27,649	7,163	50,319
<input type="checkbox"/> 7	<input checked="" type="checkbox"/>	7	1955	61,483	33,803	29,030	8,285	53,780
<input type="checkbox"/> 8	<input checked="" type="checkbox"/>	8	1956	61,483	33,803	29,030	8,285	53,780
<input type="checkbox"/> 9	<input checked="" type="checkbox"/>	9	1957	61,483	33,803	29,030	8,285	53,780
<input type="checkbox"/> 10	<input checked="" type="checkbox"/>	10	1958	61,483	33,803	29,030	8,285	53,780
<input type="checkbox"/> 11	<input checked="" type="checkbox"/>	11	1959	61,483	33,803	29,030	8,285	53,780
<input type="checkbox"/> 12	<input checked="" type="checkbox"/>	12	1960	66,227	34,517	32,178	11,658	55,636
<input type="checkbox"/> 13	<input checked="" type="checkbox"/>	13	1961	66,227	34,517	32,178	11,658	55,636
<input type="checkbox"/> 14	<input checked="" type="checkbox"/>	14	1962	66,227	34,517	32,178	11,658	55,636
<input type="checkbox"/> 15	<input checked="" type="checkbox"/>	15	1963	69,172	35,533	33,639	11,646	57,526
<input type="checkbox"/> 16	<input checked="" type="checkbox"/>	16	1964	70,499	36,142	34,357	12,950	57,549

Selected: 0 Rows: 60/73 Page size 50 Page 1 /2       

The screenshot shows a 'Delete' dialog for a CSV file named 'ChinaPopulation_en.csv'. The 'Data' tab is selected. A tooltip is displayed over the 'Data row' checkbox in the first row of the table, providing instructions for selecting rows. The table contains 73 rows of data, with rows 1 through 16 highlighted in yellow. The tooltip content includes:

- To select rows:
 - Press key SHIFT and click rows to select multiple rows of a range.
 - Press key CTRL and click rows to select multiple rows one by one.
 - Click the checkboxes in head of rows.
- To edit data:
 - Header texts of editable columns are in blue (Editable primary keys are still in red).
 - Click the editable cell, then its input field will be shown.
 - When editable cell loses focus, its value will be checked and invalid data will be rejected.
- To display/hide columns, click button "+" in right side of table header. This operation does not affect data.
- To adjust order of columns, click and drag column headers. This operation does not affect data.
- To order rows:
 - Click column header for ascending order by this column.
 - Click column header again for descending order by this column.
 - Click column header for third time to cancel ordering of this column.
- Press SHIFT and click more column headers to order by multiple columns. This operation only affect current data in table.
- Header texts of primary keys are in red and auto-increased keys are in orange.

2.11.4 Set styles / Mark Abnormal Values

2.11.4.1 Manage Styles

1. Add/Edit/Delete/Rename styles.
2. Define title and sequence number of the style.
3. Set whether the style marks abnormal values.

Set Styles / Mark Abnormal Values - CSV - 63 - d:\tmp\mybox-data-6.5.9\data\ChinaPopulation_en.csv

Set Styles / Mark Abnormal Values

Seq...	Abnormal ...	Title	From	
1.0		Analyse of Female	1	-1
2.0		Analyse of Male	1	-1
3.0	<input checked="" type="checkbox"/>	Abnormal of Fem...	25	-1
4.0	<input checked="" type="checkbox"/>	Abnormal of Male	25	-1

ID: 7      

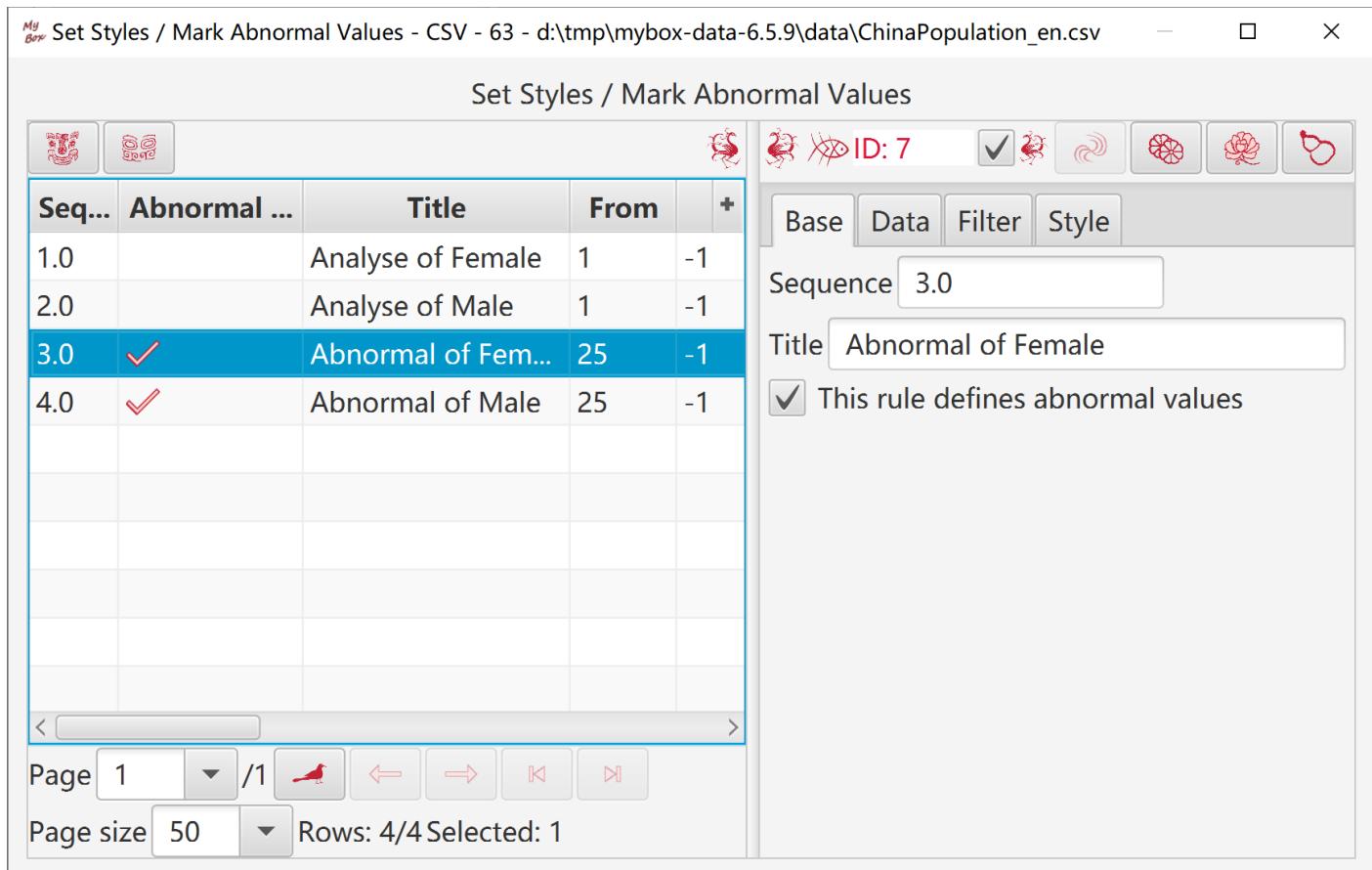
Base Data Filter Style

Sequence 3.0

Title Abnormal of Female

This rule defines abnormal values

Page 1 /1 
Page size 50 Rows: 4/4 Selected: 1



2.11.4.2 Data Scope

Define conditions to determine which data cells to apply the style:

1. Range of data rows
2. Column names.
3. Row filter.

Notice, data of a row number may be changed when some rows are added or deleted. Example, when insert 2 rows before "row 6", original "row 12" becomes "row 14" while current "row 12" was "row 10".

So "row number" is not right way to locate a specific data row while rows number is changing. A way to refer special rows is the expression composed of column values.

Seq...	Abn...	Title	From	+/-
1.0		Analyse of Female	1	-
2.0		Analyse of Male	1	-
3.0	<input checked="" type="checkbox"/>	Abnormal of Fem...	25	-
4.0	<input checked="" type="checkbox"/>	Abnormal of Male	25	-

ID: 7

Base Data Filter Style

DataRowNumber

From Blank, zero, or negative means no limitation to row number

To Blank, zero, or negative means the last row

Columns to be calculated (No selection means all)

year_ population at year-end(ten thousand)
 male(ten thousand) female(ten thousand)
 urban(ten thousand) rural(ten thousand)

< >

Page 1 /1

Page size 50 Rows: 4/4

Selected: 1

2.11.4.3 Define Styles

1. Font color, font size, background color, bold, etc.
2. More values in format of JavaFx CSS.

Set Styles / Mark Abnormal Values - CSV - 63 - d:\tmp\mybox-data-6.5.9\data\ChinaPopulation_en.csv

Set Styles / Mark Abnormal Values

Seq...	Abn...	Title	From	To
1.0		Analyse of Female	1	-1
2.0		Analyse of Male	1	-1
3.0	✓	Abnormal of Fem...	25	-1
4.0	✓	Abnormal of Male	25	-1

Page 1 /1 Rows: 4/4 Selected: 1

ID: 7

Base Data Filter Style

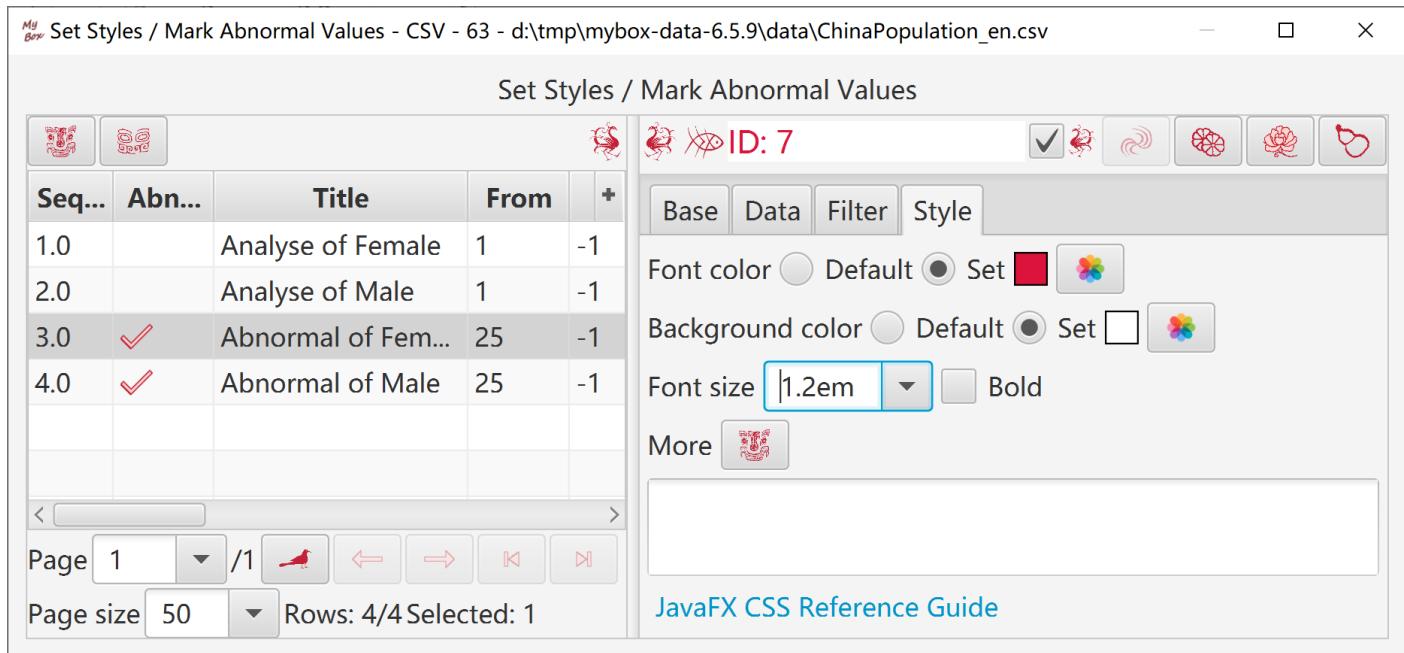
Font color Default Set 

Background color Default Set 

Font size  Bold

More 

JavaFX CSS Reference Guide



2.11.4.4 Apply Styles

When load data page, all styles of the data are applied to the rows one by one in order of their sequence numbers(from small to large).

When data scopes of styles are interlaced, style with larger number overrides the one with smaller number for the interlaced part.

The screenshot shows a data grid with the following approximate data:

Table	Data	year_	populatio...	male(ten t...)	female(te...)	urban(ten ...)	rural(ten t...)
24	24	1972	87,177	44,813	42,364	14,935	72,242
25	25	1973	89,211	45,876	43,335	15,345	73,866
26	26	1974	90,859	46,727	44,132	15,595	75,264
27	27	1975	92,420	47,564	44,856	16,030	76,390
28	28	1976	93,717	48,257	45,460	16,341	77,376
29	29	1977	94,974	48,908	46,066	16,669	78,305
30	30	1978	96,259	49,567	46,692	17,245	79,014
31	31	1979	97,542	50,192	47,350	18,495	79,047
32	32	1980	98,705	50,785	47,920	19,140	79,565
33	33	1981	100,072	51,519	48,553	20,171	79,901
34	34	1982	101,654	52,352	49,302	21,480	80,174
35	35	1983	103,008	53,152	49,856	22,274	80,734
36	36	1984	104,357	53,848	50,509	24,017	80,340

Selected: 0 Rows: 50/73 Page size 50 Page 1 /2

2.11.5 Paste Content in System Clipboard

1. Read contents in System Clipboard.
2. Guess delimiter and parse data.
3. User can select a delimiter to parse data.
4. Select how to parse:
 - CSV parser: Values can contain delimiter and line break. But not support parsing “Blanks” and “Regular Expression”.
 - Text parser: Values can not contain delimiter and line break. Support parsing “Blanks” and “Regular Expression”.
5. Select whether first line defines column names.
6. Based on parsed data, select rows and columns to be pasted.
7. Select location to paste data: row and column in target data.
8. Select the way to paste data: Replace, Insert Above, or Append Below.

Paste content in System Clipboard : CSV - 957 - d:\tmp\mybox-data-6.5.9\AppData\Temp\2022-11-14_15-42-45-08...

Table r...	Data...	C...	Co...	Colu...
<input type="checkbox"/> 1	1	age	sex	BMI(bod
<input checked="" type="checkbox"/> 2	2	59	2	32.1
<input checked="" type="checkbox"/> 3	3	48	1	21.6
<input type="checkbox"/> 4	4	72	2	30.5
<input checked="" type="checkbox"/> 5	5	24	1	25.3
<input checked="" type="checkbox"/> 6	6	50	1	23
<input type="checkbox"/> 7	7	23	1	22.6
<input checked="" type="checkbox"/> 8	8	36	2	22
<input type="checkbox"/> 9	9	66	2	26.2
<input type="checkbox"/> 10	10	60	2	32.1

If a value contains delimiter or line break, it should be surrounded by quotes.

2.11.6 Paste Content in MyBox Clipboard

1. Select data item in MyBox Clipboard.
2. Select rows and columns to be pasted.
3. Select location to paste data: row and column in target data.
4. Select the way to paste data: Replace, Insert Above, or Append Below.

Paste content in MyBox Clipboard : MyBox Clipboard - 1089 - ChinaPopulation_zh.csv

Table row	ID	Type	Name
<input checked="" type="checkbox"/> 1	1089	MyBox ...	ChinaPopu
<input type="checkbox"/> 2	1086	MyBox ...	ChinaPopu
<input type="checkbox"/> 3	1085	MyBox ...	ChinaPopu
<input type="checkbox"/> 4	1083	MyBox ...	ChinaPopu
<input type="checkbox"/> 5	1082	MyBox ...	ChinaPopu
<input type="checkbox"/> 6	1063	MyBox ...	a
<input type="checkbox"/> 7	6	MyBox ...	b

Page 1 /1 Rows: 7/7 Selected: 1

Table row	Data row	源行号	年	年末总人
	1	1	年	年末总人
✓	2	1	1949	54,167
✓	3	2	1950	55,196
✓	4	3	1951	56,300
	5	4	1952	57,482
	6	5	1953	58,796

Location to paste
Table row 1
Column year_
Replace

2.12 Trim Data

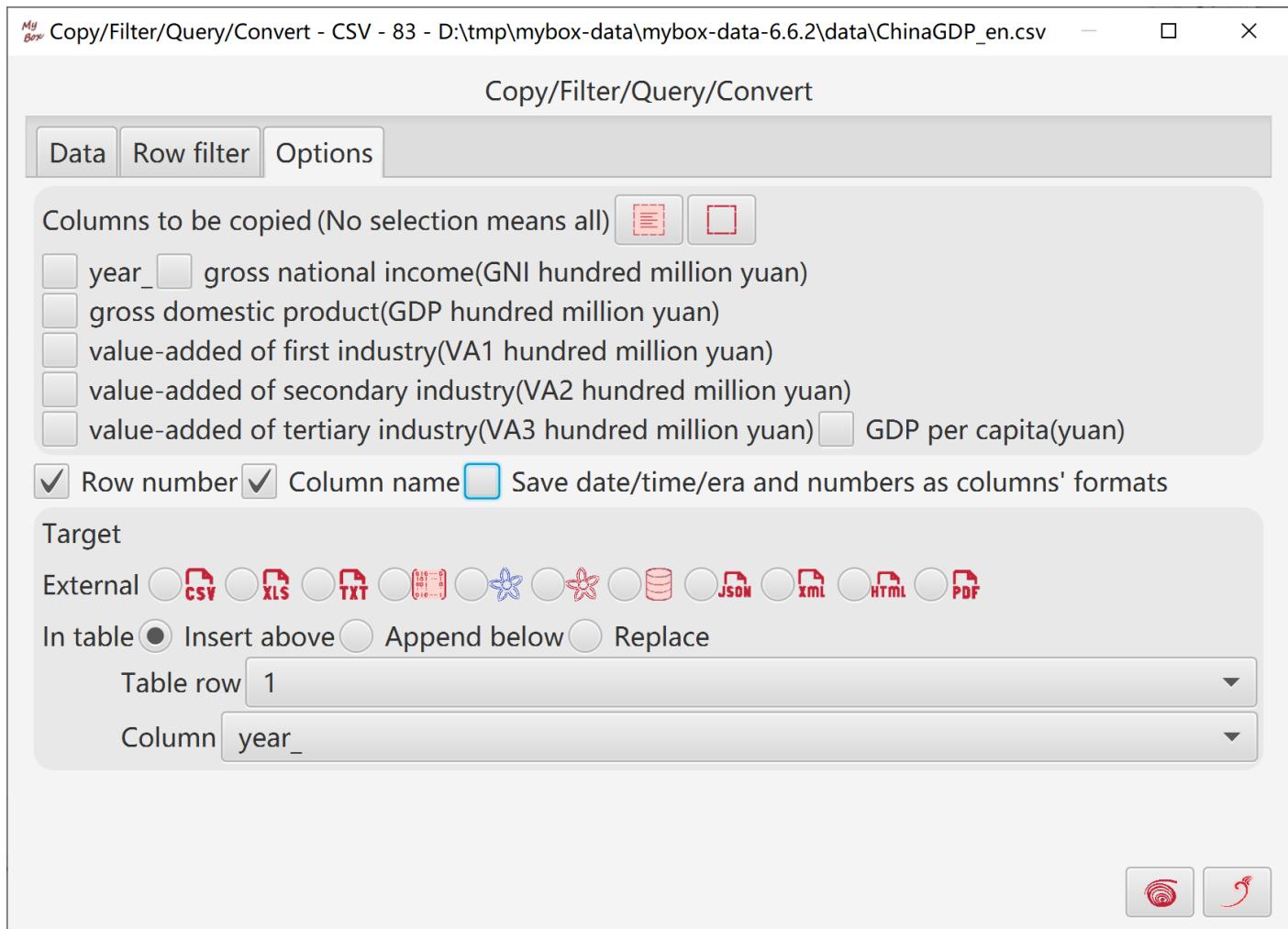
Hover or click button “Function” and select functions under menu item “Trim”.

The screenshot shows the MyBox Data Tools interface with a CSV file named "ChinaPopulation_en.csv" open. A context menu is displayed over a table selection, specifically over the first row. The menu path is "Data" > "Modify" > "Trim". The "Trim" option is highlighted with a blue background. Other options in the "Modify" submenu include "Copy/Filter/Query/Convert", "Sort", "Transpose", "Normalize" (which is also highlighted with a blue background), and "Split/Group". The main window displays a table with columns: year_, populatio..., male(ten t...), fe... (partially visible). The table contains 15 rows of data from 1949 to 1963. At the bottom of the window, there are status bars for "Selected: 0", "Rows: [1-50]50/73", "Page size 50", "Page 1 /2", and navigation icons.

Table r...	Data...	year_	populatio...	male(ten t...)	fe...
1	1	1949			
2	2	1950			
3	3	1951			
4	4	1952			
5	5	1953			
6	6	1954			
7	7	1955	61465	31809	29
8	8	1956	62828	32536	30
9	9	1957	64653	33469	31 104
10	10	1958	65994	34195	31799
11	11	1959	67207	34890	32317
12	12	1960	66207	34283	31924
13	13	1961	65859	33880	31979
14	14	1962	67296	34517	32778
15	15	1963	69172	35533	33639
...

2.12.1 Copy/Filter/Query/Convert

1. Select data rows. Row filter can be set.
2. Select columns to be copied.
3. Select whether:
 - Copy row numbers
 - Copy column names
 - Save date/time/era and numbers as columns' formats
4. Results can be written as external data, including: new csv/excel/text file, matrix, system clipboard, MyBox clipboard, database table, json, xml, html, or pdf.
5. When handle rows for current page or selected ones, results can be written in current page: insert, append, or replace data in defined location of the table.



2.12.2 Sort

1. Select data rows. Row filter can be set.
2. Select columns to sort and whether descending.
Data type of column affects sorting results.
3. Maximum rows number of results can be set.
4. Select columns to be copied.
5. Select whether copy row number or column name.
6. Results can be written as external data, including: new csv/excel/text file, matrix, system clipboard, MyBox clipboard, database table, json, xml, html, or pdf.
7. When handle rows for current page or selected ones, results can be written in current page: insert, append, or replace data in defined location of the table.

Sort - CSV - 83 - D:\tmp\mybox-data\mybox-data-6.6.2\data\ChinaGDP_en.csv

Sort

Data Row filter Options

Order by(Column type affects sorting results)

<input type="checkbox"/> Table r...	Sort		
<input type="checkbox"/> 1	year_-Descending		
<input type="checkbox"/> 2	year_-Ascending		
<input checked="" type="checkbox"/> 3	gross national income(GNI hundred million yuan)-Descending		
<input type="checkbox"/> 4	gross national income(GNI hundred million yuan)-Ascending		
<input type="checkbox"/> 5	gross domestic product(GDP hundred million yuan)-Descending		

Maximum sorted rows to take(Empty/zero/negative to unlimit)

Columns to be copied (No selection means all)

year_ gross national income(GNI hundred million yuan)
 gross domestic product(GDP hundred million yuan)
 value-added of first industry(VA1 hundred million yuan)
 value-added of secondary industry(VA2 hundred million yuan)
 value-added of tertiary industry(VA3 hundred million yuan) GDP per capita(yuan)

Row number Column name

Target

External

Name

In table Insert above Append below Replace

2.12.3 Transpose

1. Select data rows. Row filter can be set.
2. Select columns to be calculated.
3. Select whether:
 - Copy row number
 - Copy column name
 - Set first column as column names.
4. Results can be written as external data, including: new csv/excel/text file, matrix, system clipboard, MyBox clipboard, database table, json, xml, html, or pdf.
5. When handle rows for current page or selected ones, results can be written in current page: insert, append, or replace data in defined location of the table.

Transpose - CSV - 83 - D:\tmp\mybox-data\mybox-data-6.6.2\data\ChinaGDP_en.csv

Transpose

Data Row filter Options

Columns to be calculated (No selection means all)

year_ gross national income(GNI hundred million yuan)
 gross domestic product(GDP hundred million yuan)
 value-added of first industry(VA1 hundred million yuan)
 value-added of secondary industry(VA2 hundred million yuan)
 value-added of tertiary industry(VA3 hundred million yuan) GDP per capita(yuan)

Row number Column name First column as column names

Target

External CSV XLS TXT Matrix JSON XML HTML PDF

Name: Gross domestic product(GDP) of China

In table Insert above Append below Replace

2.12.4 Normalization

1. Select data rows. Row filter can be set.
2. Select columns to be calculated.
3. Select objects: According to Columns/rows/all.
4. Select algorithms: MinMax(Range can be set), sum(L1), Zscore(L2).
5. To non-numeric, skip or count as zero.
6. Select columns to be copied.
7. Results can be written as external data, including: new csv/excel/text file, matrix, system clipboard, MyBox clipboard, database table, json, xml, html, or pdf.
8. When handle rows for current page or selected ones, results can be written in current page: insert, append, or replace data in defined location of the table.

Normalize - CSV - 84 - D:\tmp\mybox-data\mybox-data-6.6.2\data\IrisSpecies_en.csv

Normalize

Data **Row filter** **Options**

Columns to be calculated (No selection means all) SepalLengthCm SepalWidthCm PetalLengthCm PetalWidthCm Species

According to Columns Rows All

Algorithm MinMax L1(Sum) L2(Z-Score)

Range , (-1,1) [0.0,1.0]

Row number Column name

Decimal scale

To invalid number Skip Count as zero

Columns to be copied Species

SepalLengthCm SepalWidthCm PetalLengthCm PetalWidthCm Species

Target

External

Name Iris species

In table Insert above Append below Replace

2.12.5 Split/Group

1. Select data rows. Row filter can be set.
2. Set parameters of grouping.
3. Select columns to be copied.
4. Set sort conditions and maximum data rows in each group.
5. Select whether include row numbers.
6. Select target:
 - All groups in single file
 - One file for each group
 - Database table

MyBox Split/Group - CSV - 958 - d:\tmp\mybox-data-6.5.9\data\ChinaGDP_en.csv

Split/Group

Data **Row filter** **Group** **Options**

Columns to be copied (No selection means all)

year_ gross national income(GNI hundred million yuan)
 gross domestic product(GDP hundred million yuan)
 value-added of first industry(VA1 hundred million yuan)
 value-added of secondary industry(VA2 hundred million yuan)
 value-added of tertiary industry(VA3 hundred million yuan) GDP per capita(yuan)

Order by(Column type affects sorting results)

Table r...	Sort	+
<input type="checkbox"/> 1	year_-Descending	
<input type="checkbox"/> 2	year_-Ascending	
<input checked="" type="checkbox"/> 3	gross national income(GNI hundred million yuan)-Descending	
<input type="checkbox"/> 4	gross national income(GNI hundred million yuan)-Ascending	
<input type="checkbox"/> 5	gross domestic product(GDP hundred million yuan)-Descending	

Maximum rows in each group(Empty/zero/negative to unlimit)

Row number

Target All groups in single file One file for each group Database Table

2.13 Data Calculation

Hover or click button “Function” to select functions under menu item “Calculation”.

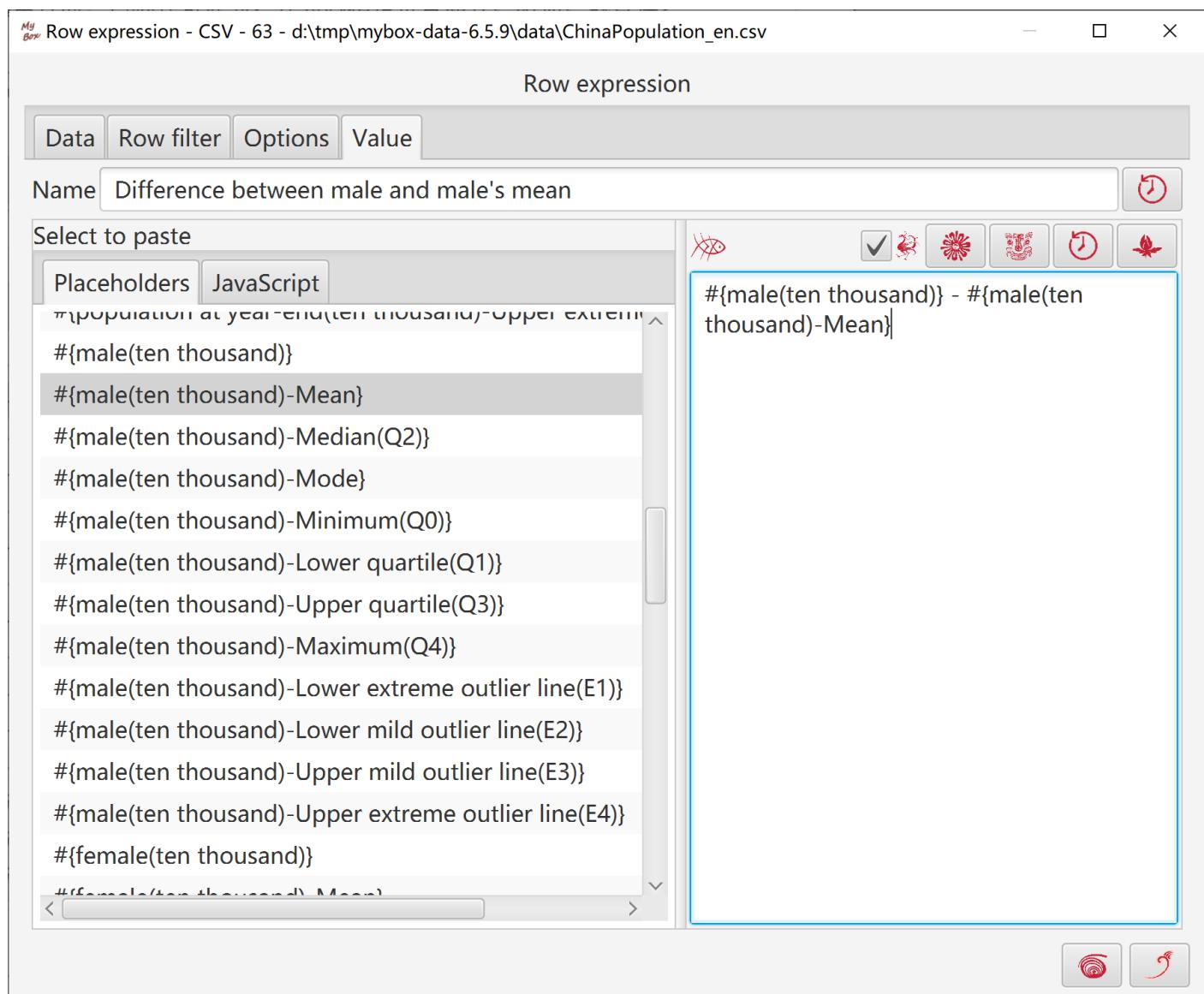
The screenshot shows a CSV file titled "Edit CSV File : CSV - 958 - d:\tmp\mybox-data-6.5.9\data\ChinaGDP_en.csv" in the MyBox interface. The menu bar includes Window, Document, Image, Data, File, Media, Network, Settings, Recent Accessed, Development, and Help. Below the menu bar are several icons. The main area displays a table of GDP data for China from 1952 to 1967. A context menu is open over the table, with "Calculation" selected. Other options in the menu include Data, Modify, Trim, Charts, Group Charts, Examples, and a checkbox for "Pop when mouse hovering". At the bottom of the interface, there are buttons for Selected: 0, Rows: [1-50]50/70, Page size: 50, Page: 1, and navigation arrows.

Table r...	Data...	year_	gross nati...	gross do...	v
1	1	1952	679.1	679.1	34
2	2	1953	679.1	679.1	34
3	3	1954	679.1	679.1	34
4	4	1955	679.1	679.1	34
5	5	1956	679.1	679.1	34
6	6	1957	679.1	679.1	34
7	7	1958	679.1	679.1	34
8	8	1959	679.1	679.1	34
9	9	1960	679.1	679.1	34
10	10	1961	679.1	679.1	34
11	11	1962	679.1	679.1	34
12	12	1963	1248.5	1248.5	47.5
13	13	1964	1469.9	1469.9	559
14	14	1965	1734	1734	651.1
15	15	1966	1888.7	1888.7	702.2
16	16	1967	1704.2	1704.2	711.2

Selected: 0 Rows: [1-50]50/70 Page size: 50 Page: 1 /2

2.13.1 Row Expression

1. Select data row. Data filter can be set.
2. Select columns to be copied.
3. Input row expression and its name.
4. Results can be written as external data, including: new csv/excel/text file, matrix, system clipboard, MyBox clipboard, database table, json, xml, html, or pdf.
5. When handle rows for current page or selected ones, results can be written in current page: insert, append, or replace data in defined location of the table.



2.13.2 Descriptive Statistic

1. Select data rows. Data filter can be set.
2. Select columns to be calculated.
3. Select statistic items:
 count, sum, mean, geometric mean, sum of squares, mode, skewness, population variance, sample variance, population standard deviation, sample standard deviation, minimum(Q0), lower quartile(Q1), median(Q2), upper quartile(Q3), maximum(Q4), upper extreme outlier line(E4), upper mild outlier line(E3), lower mild outlier line(E2), lower extreme outlier line(E1)
4. Select objects: According to columns/rows/all.
5. Set decimal scale.
6. To non-numeric: skip or count as zero
7. Results can be written as external data, including: new csv/excel/text file, matrix, system clipboard, MyBox clipboard, database table, json, xml, html, or pdf.
8. When handle rows for current page or selected ones, results can be written in current page: insert, append, or replace data in defined location of the table.

MyBox Descriptive statistics - CSV - 84 - D:\tmp\mybox-data\mybox-data-6.6.2\data\IrisSpecies_en.csv

Descriptive statistics

Data Row filter Options

Columns to be calculated (No selection means all)

SepalLengthCm SepalWidthCm PetalLengthCm PetalWidthCm Species

Descriptive statistics

Count Summation Mean Geometric mean Sum of squares
 Population variance Sample variance Population standard deviation Sample standard deviation
 Skewness
 Minimum(Q0) Lower quartile(Q1) Median(Q2) Upper quartile(Q3) Maximum(Q4)
 Upper extreme outlier line(E4) Upper mild outlier line(E3) Lower mild outlier line(E2)
 Lower extreme outlier line(E1) Mode

According to Columns Rows All

Decimal scale 2

To invalid number Skip Count as zero

Target

External CSV XLS TXT JSON XML HTML PDF

Name Iris species

In table Insert above Append below Replace

2.13.3 Statistic in Groups

2.13.3.1 Options of Calculation

1. Select data rows. Data filter can be set.
2. Set parameters of grouping.
3. Select columns to be calculated.
4. Select statistic items. “Count” is always selected.
5. Set decimal scale.
6. To invalid numbers: set as blank or zero.

Statistic in groups - CSV - 958 - d:\tmp\mybox-data-6.5.9\data\ChinaGDP_en.csv

Data **Row filter** **Group** **Options**

Columns to be calculated (No selection means all)

year_ gross national income(GNI hundred million yuan)
 gross domestic product(GDP hundred million yuan)
 value-added of first industry(VA1 hundred million yuan)
 value-added of secondary industry(VA2 hundred million yuan)
 value-added of tertiary industry(VA3 hundred million yuan)
 GDP per capita(yuan)

Descriptive statistics

Count Summation Mean Geometric mean
 Sum of squares
 Population variance Sample variance
 Population standard deviation Sample standard deviation
 Skewness
 Minimum(Q0) Lower quartile(Q1) Median(Q2)
 Upper quartile(Q3) Maximum(Q4)
 Upper extreme outlier line(E4) Upper mild outlier line(E3)
 Lower mild outlier line(E2) Lower extreme outlier line(E1)
 Mode

Decimal scale

To invalid number Set as blank Set as zero

Display all data in chart

Group Data **Statistic data** **Chart data**

Data...	id	group_id	range_year_
1	1	1	[1952,1961)
2	2	1	[1952,1961)
3	3	1	[1952,1961)
4	4	1	[1952,1961)
5	5	1	[1952,1961)
6	6	1	[1952,1961)
7	7	1	[1952,1961)
8	8	1	[1952,1961)
9	9	1	[1952,1961)
10	10	2	[1961,1970)
11	11	2	[1961,1970)
12	12	2	[1961,1970)
13	13	2	[1961,1970)
14	14	2	[1961,1970)
15	15	2	[1961,1970)
16	16	2	[1961,1970)

Selected: 0 Rows: [1-20]20/70 Page size
 Page /4

2.13.3.2 Grouped Data

After data are grouped, 2 columns are added: “Group id”, “Group Parameter”.

Statistic in groups - CSV - 958 - d:\tmp\mybox-data-6.5.9\data\ChinaGDP_en.csv

Data...	id	group_id	range_year_	year_	gross_nati...	gross_do...	value_add...	value_a...
1	1	1	[1952,1961)	1952	679.1	679.1	342.9	141.1
2	2	1	[1952,1961)	1953	824.4	824.4	378	191.6
3	3	1	[1952,1961)	1954	859.8	859.8	392	210.8
4	4	1	[1952,1961)	1955	911.6	911.6	421	221.5
5	5	1	[1952,1961)	1956	1030.7	1030.7	443.9	280.4
6	6	1	[1952,1961)	1957	1071.4	1071.4	430	316.6
7	7	1	[1952,1961)	1958	1312.3	1312.3	445.9	483.6
8	8	1	[1952,1961)	1959	1447.5	1447.5	383.8	616.7
9	9	1	[1952,1961)	1960	1470.1	1470.1	340.7	652.6
10	10	2	[1961,1970)	1961	1232.3	1232.3	441.1	393.5

Selected: 0 Rows: [1-20]20/70 Page size 20 Page 1 /4

2.13.3.3 Statistic Data

Based on grouped data, calculate statistic values for each group.

Statistic in groups - CSV - 958 - d:\tmp\mybox-data-6.5.9\data\ChinaGDP_en.csv

Data...	id	group_id	range_year_	column_name	group_count	group_sum	group_mean	grc +
1	1	1	[1952,1961)	gross national i...	9	9606.9	1067.43	103
2	2	1	[1952,1961)	gross domestic...	9	9606.9	1067.43	103
3	3	1	[1952,1961)	value-added of...	9	3578.2	397.58	395
4	4	1	[1952,1961)	value-added of...	9	3114.9	346.1	303
5	5	1	[1952,1961)	value-added of...	9	2913.9	323.77	311
6	6	1	[1952,1961)	GDP per capita...	9	1527	169.67	166
7	7	2	[1961,1970)	gross national i...	9	14235.9	1581.77	155
8	8	2	[1961,1970)	gross domestic...	9	14235.9	1581.77	155
9	9	2	[1961,1970)	value-added of...	9	5480.7	608.97	597

Selected: 0 Rows: [1-20]20/48 Page size 20 Page 1 /3

2.13.3.4 Data in Chart

1. Statistic data are trimmed into columns and rows for chart.
2. If select “Display all data in chart”, then data of all pages are displayed in chart. Or else only data of current page are displayed in the chart.

Statistic in groups - CSV - 958 - d:\tmp\mybox-data-6.5.9\data\ChinaGDP_en.csv

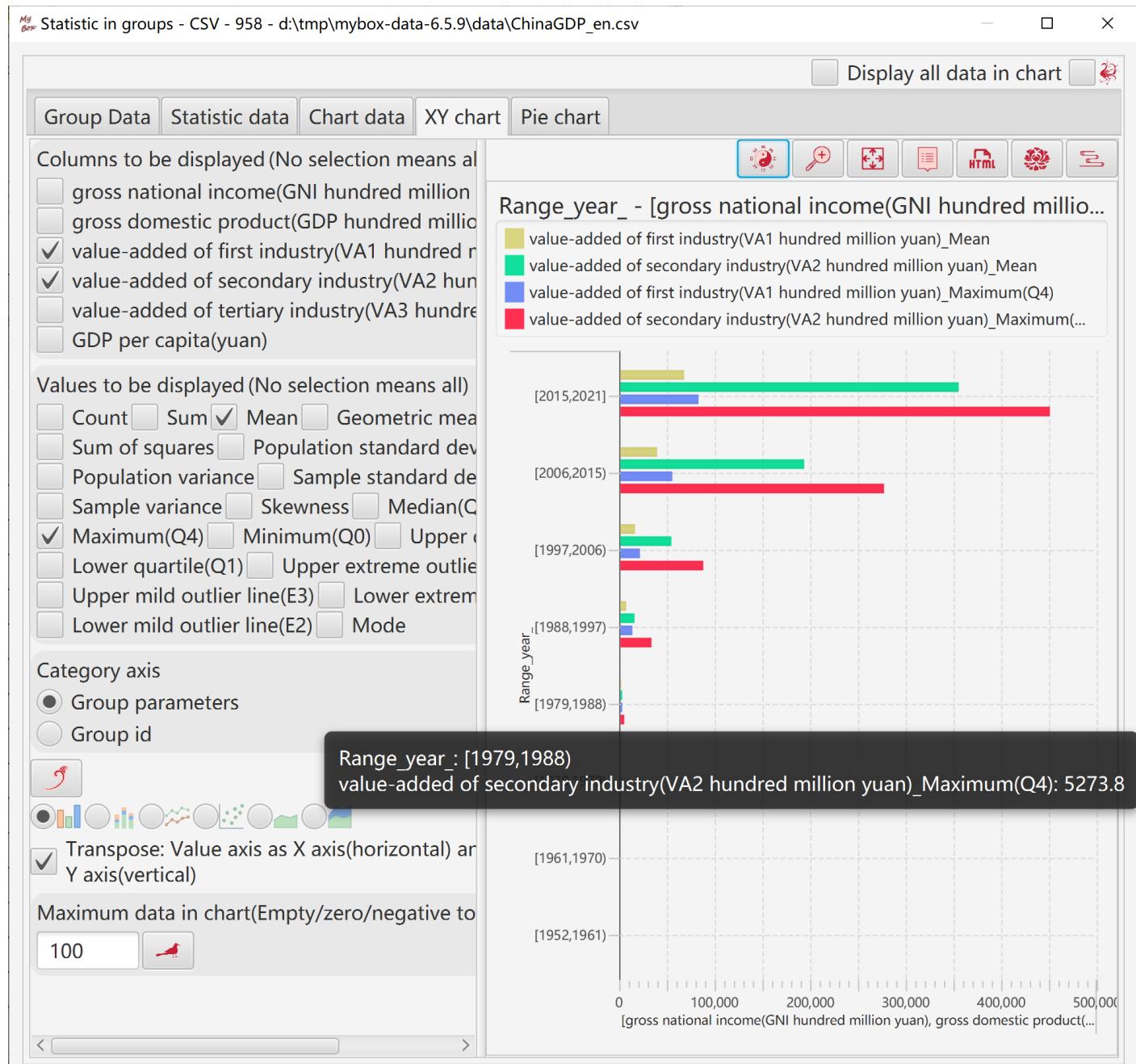
Data...	Group id	Range_year_	Group_Count	gross national income(GNI hundred million yuan)_Sum	gross i +
1	1	[1952,1961)	9	9606.9	1067.43
2	2	[1961,1970)	9	14235.9	1581.77
3	3	[1970,1979)	9	25829.7	2869.97
4	4	[1979,1988)	9	64024.4	7113.82
5	5	[1988,1997)	9	315828.2	35092.02
6	6	[1997,2006)	9	1064800.3	118311.14
7	7	[2006,2015)	9	3822494.2	424721.58
8	8	[2015,2021]	7	6296896.8	899556.69

Selected: 0 Rows: [1-8]8/8 Page size 20 Page 1 /1

2.13.3.5 XY Chart

Statistic data of groups can be displayed as XY chart:

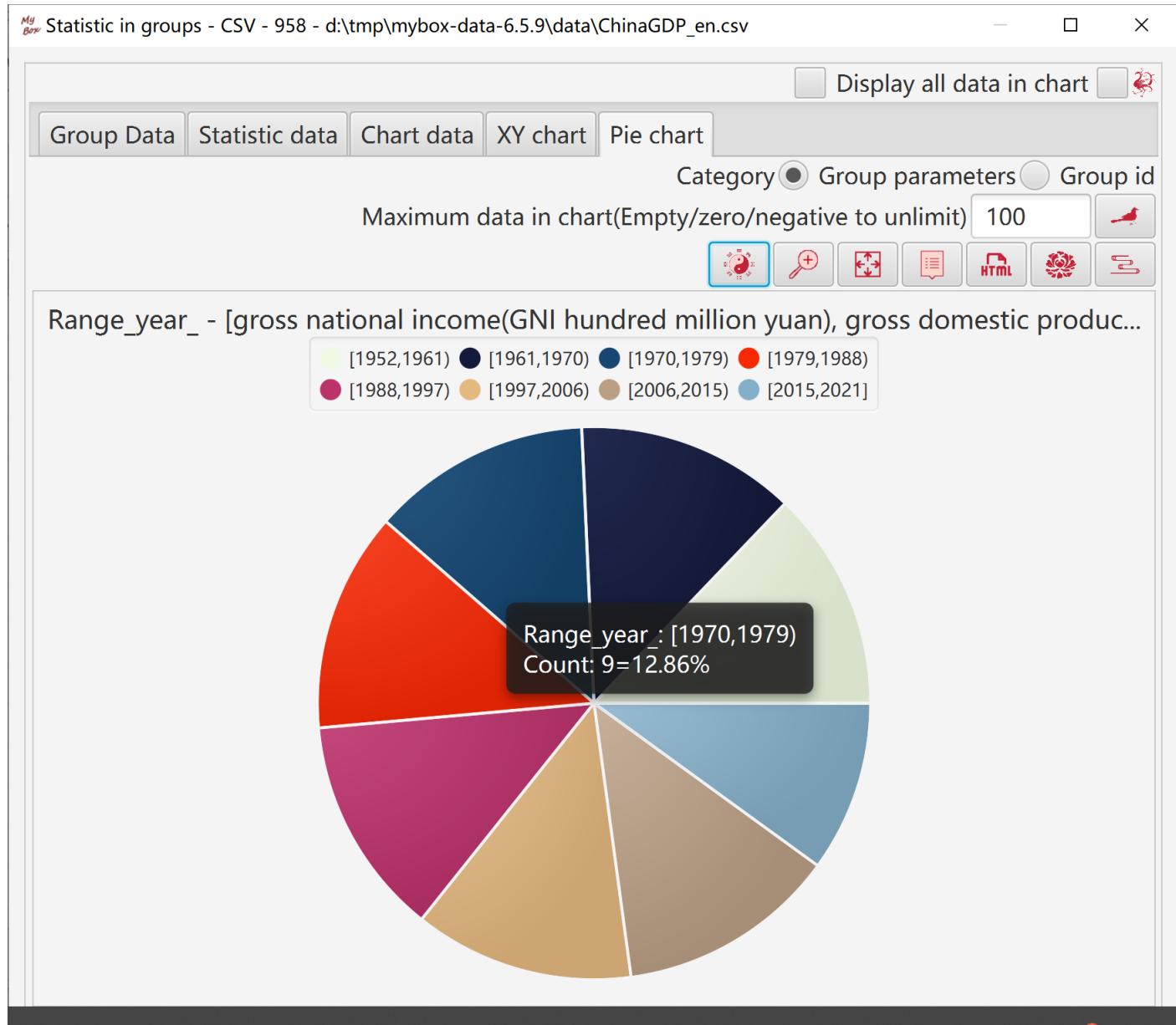
1. Select columns to be displayed
2. Select statistic values to be displayed.
3. Select category axis: group parameter, or group id.
4. Select type of XY chart.
5. Select whether transpose the chart.
6. Set maximum data number in the chart.



2.13.3.6 Pie Chart

Groups and their “Count” are displayed as Pie chart.

1. Select category axis: group parameter, or group id.
2. Set maximum data number in the chart.



2.13.4 Simple Linear Regression

2.13.4.1 Base of Implementation

This tool is based on Apache Commons Math.

The regression does not store data, so calculation itself has not memory limitation when handle lots of rows.

2.13.4.2 Options of Calculation

1. Select data rows. Data filter can be set.
2. Select one column as independent variable.
3. Select another column as dependent variable. It should not be the same column of independent variable.
4. Set desired significance level(alpha).
5. Select whether includes intercept.
6. Select whether display all data in chart.
7. Set decimal scale.

2.13.4.3 Regression

Display values status of regression steps in table, including number of observations, slope, intercept, coefficient of determination(R-Square), correlation coefficient(R), mean of squared error(MSE) , sum of squared errors(SSE), total sum of squares(SSTO), sum of squares about regression(SSM/SSR), etc.

N...	income	happiness	Slope(b1)	Intercept(b0)	Coeffici...	Pearson
1	3.8626	2.3145	NaN	NaN	NaN	NaN
2	4.6399	3.7379	1.8314	-4.7594	1	1
3	2.1347	0.2687	1.3514	-2.6845	0.9874	0.9937
4	6.5013	4.3748	0.9625	-1.45	0.9171	0.9577
5	3.6512	2.1558	0.9579	-1.4127	0.9183	0.9583
6	2.2865	1.8936	0.8249	-0.715	0.8447	0.9191
7	4.7489	4.903	0.9204	-0.8517	0.7453	0.8633
8	5.4592	4.8335	0.975	-0.9964	0.7754	0.8806

2.13.4.4 Model

1. Display fitted linear model.
2. Display data status of last regression step.
3. Input value for independent variable, and generate predicted value.

Simple linear regression - CSV - 784 - d:\tmp\mybox-data-6.5.9\data\IncomeHappiness_en.csv

About linear regression Decimal scale 4

Model Regression Fitting Residual

Linear model: happiness = 0.2222 + 0.7097 * income

Independent variable: income = Predict

Dependent variable: happiness =

Last status

Name	Value
Row number	477
Number of observations	477
income	4.4981
happiness	1.9071
Slope(b1)	0.7097
Intercept(b0)	0.2222
Coefficient of determination(R-Square)	0.7517
Pearson's product moment correlation coefficient(R)	0.867
Mean of squared error(MSE)	0.5057

DIV

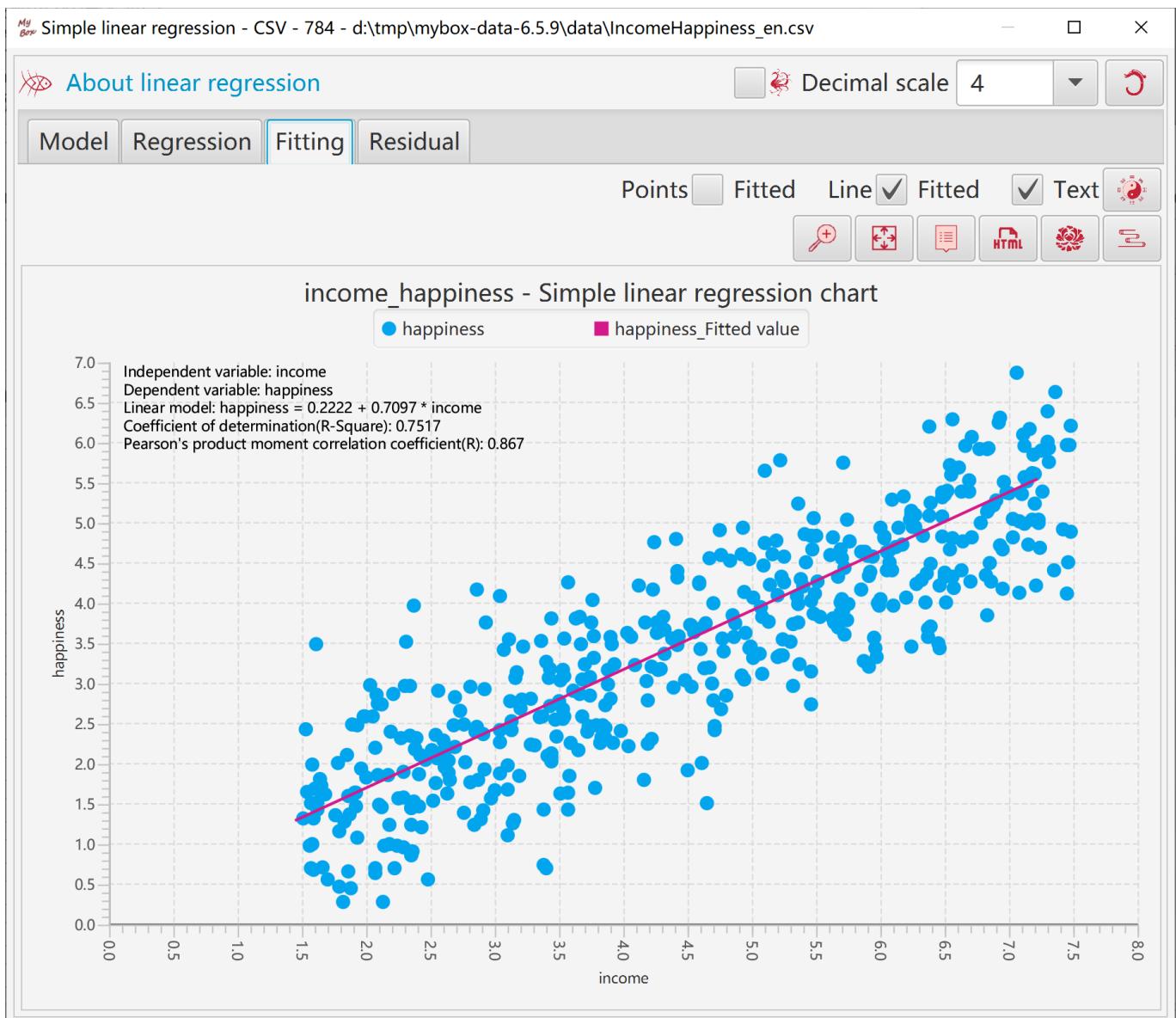
2.13.4.5 Fitted Chart

1. When handle all data rows(all pages), option to display all values in chart.

When display all values in chart, need concern memory limitation when load lots of data.

Or else only values in current page will be displayed in chart while all pages involve regression, and no memory limitation.

2. Set parameters of plot, X axis and Y axis.
3. Options to display fitted points, fitted line, or model description in chart.
4. Options to display data labels.
5. Set random colors to fitted points/line.
6. Fitted chart can be popped.
7. Html including fitted chart and its data can be created.
8. Display fitted chart's data in table.



2.13.4.6 Residual Chart

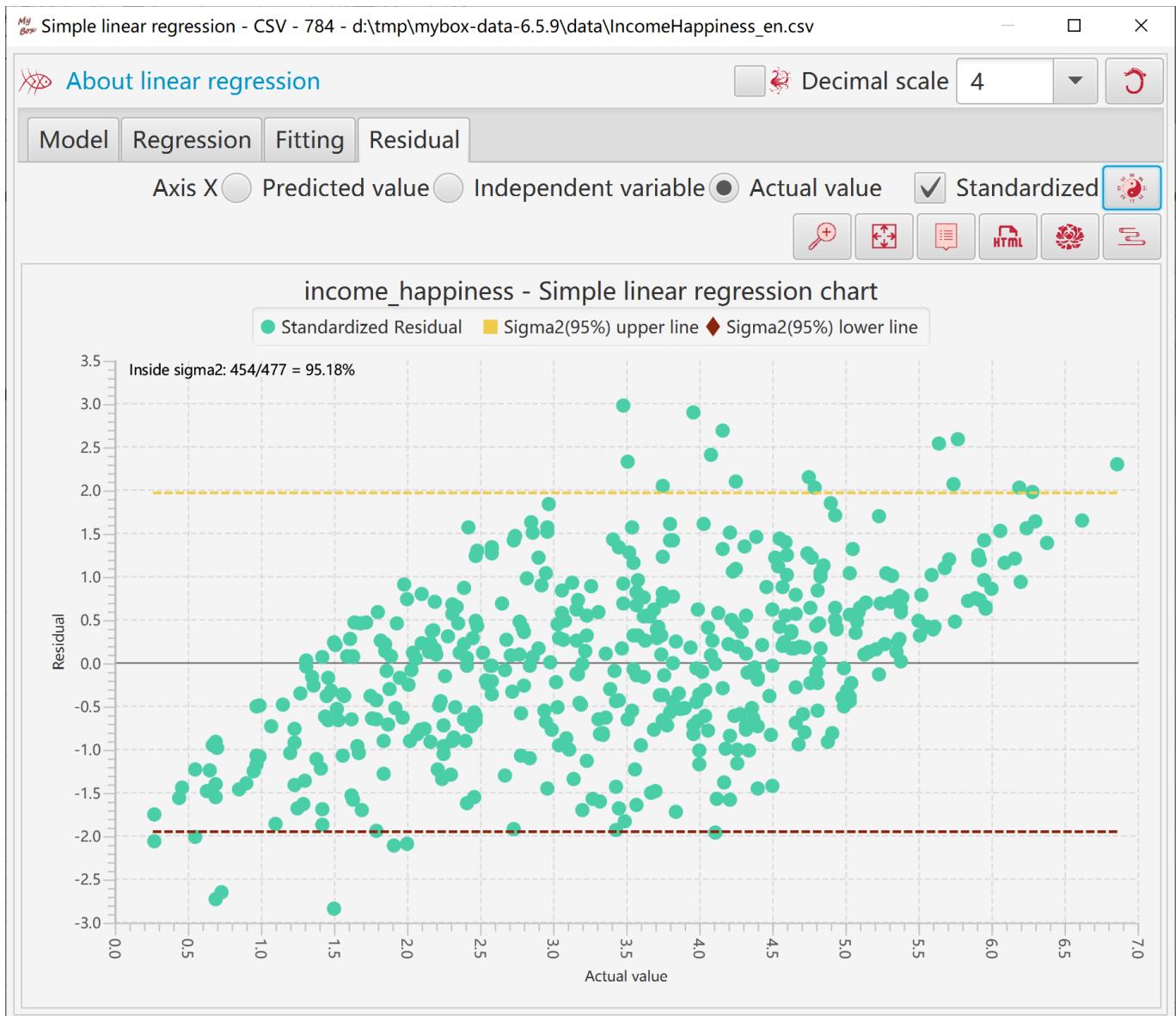
1. X axis can be set as: predicted value, independent variable, or actual value.

2. Option whether standardize residuals.

When standardize residuals, upper line and lower line of Sigma2(95%) will be displayed.

3. Set random color to points and lines.

4. Display residual chart's data in table.



2.13.5 Simple Linear Regression - Combination

This tool helps to generate data of simple linear regression:

1. Select some columns as candidates of independent variables.
2. Select some columns as candidates of dependent variables.
3. Select options like decimal scale, alpha, whether include intercept.
4. When click button "OK":

- Make pairs from candidates : one as independent variable and another as dependent variable.
 - Calculate the regression models.
 - Sort the modes by their coefficient of determination(R-Square) in descending order.
5. Select one model and click button "View" to view its regression data, fitting chart, and residual chart.

The screenshot shows the MyBox Data Tools interface for a simple linear regression analysis. The left panel contains two sections: 'Candidates of independent variables' and 'Candidates of dependent variables'. In the independent variables section, several checkboxes are checked: age, sex, BMI, BP, S1, S2, S3, S4, S5, and S6. In the dependent variables section, only 'disease progression one year after baseline' is checked. Below these sections are settings for 'Decimal scale' (set to 8) and 'Desired significance level(alpha)' (set to 0.05), along with a checkbox for 'Intercept(b0)'. The right panel is titled 'About data analysis' and contains a table with 12 rows of data. The columns are labeled 'Data...', 'Dependen...', 'Independ...', 'Coeffi...', 'Pearso...', and 'Model'. The first row has the highest coefficient value of 0.34392... and the highest Pearson correlation of 0.58645..., both for the relationship between 'disease pr...' and 'BMI(body ...)'. The table also includes other rows for various independent variables like S5, S4, S3, S2, S1, and S6, and their correlations with the dependent variable.

Data...	Dependen...	Independ...	Coeffi...	Pearso...	Model
3	disease pr...	BMI(body ...)	0.34392...	0.58645...	disease progression one year
9	disease pr...	S5(blood s...	0.32022...	0.56588...	disease progression one year
4	disease pr...	BP(averag...	0.19490...	0.44148...	disease progression one year
8	disease pr...	S4(blood s...	0.18528...	0.43045...	disease progression one year
7	disease pr...	S3(blood s...	0.15585...	0.39478...	disease progression one year
10	disease pr...	S6(blood s...	0.14629...	0.38248...	disease progression one year
5	disease pr...	S1(blood s...	0.04495...	0.21202...	disease progression one year
1	disease pr...	age	0.03530...	0.18788...	disease progression one year
6	disease pr...	S2(blood s...	0.03029...	0.17405...	disease progression one year
2	disease pr...	sex	0.00185...	0.043062	disease progression one year

2.13.6 Multiple Linear Regression

2.13.6.1 Regression

This tool helps to generate data of multiple linear regression based on Apache Commons Math:

1. Select some column as independent variables, whose data should be numbers.
2. Select a column as dependent variable, whose data should be numbers.
3. Select whether include intercept.

4. Click button "OK":

- Tool normalizes involved data by Z-Score.
- Tool calculates multiple linear regression by Ordinary Least Squares(OLS).
- Results include intercept, coefficients, R-Square, adjusted R-Square.

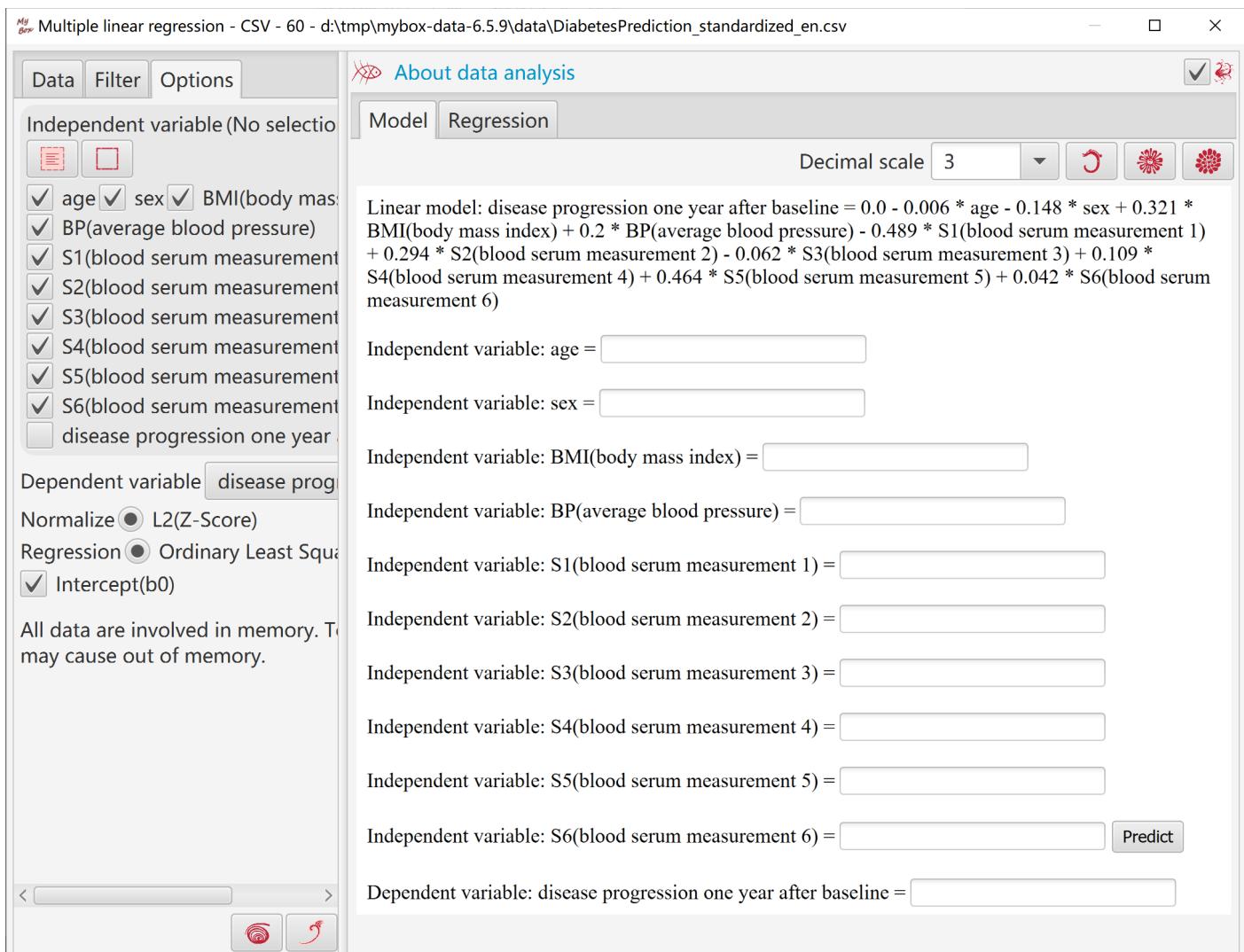
The screenshot shows the 'About data analysis' window from the MyBox Data Tools interface. The window has tabs for 'Model' and 'Regression'. The 'Model' tab is selected, displaying a table of model statistics. The table has two columns: 'Name' and 'Value'.

Name	Value
Dependent variable	disease progression one year after baseline
Independent variable	[age, sex, BMI(body mass index), BP(average blood pressure), S1(blood serum measurement 1), S2(blood serum measurement 2), S3(blood serum measurement 3), S4(blood serum measurement 4), S5(blood serum measurement 5), S6(blood serum measurement 6)]
Number of observations	442
Intercept(b0)	0.0
Coefficients	[-0.006, -0.148, 0.321, 0.2, -0.489, 0.294, -0.062, 0.109, 0.464, 0.042]
Coefficient of determination(R-Square)	0.518
Adjusted R squared	0.507
Standard Error	0.702
Variance	1.0

The left panel of the window shows configuration for the regression model. It includes sections for 'Independent variable (No selection means all)', 'Dependent variable' (set to 'disease progression one year after baseline'), 'Normalize' (set to 'L2(Z-Score)'), 'Regression' (set to 'Ordinary Least Squares(OLS)'), and a checkbox for 'Intercept(b0)'. A note at the bottom states: 'All data are involved in memory. Too many data may cause out of memory.'

2.13.6.2 Model

User can input values of independent variables and predict value of dependent variable.



2.13.7 Multiple Linear Regression - Combination

This tool helps to generate data of multiple linear regression based on Apache Commons Math:

1. Select some columns as candidates of independent variables.
2. Select some columns as candidates of dependent variables.

3. Select whether include intercept.
4. When click button "OK", the tool does following:
 - Make combination of candidates as independent variables and dependent variable .
 - Calculate the regression models.
 - Sort the modes by their adjusted coefficient of determination(R-Square) in descending order.
5. Select one model and click button "View" to view its regression data

Multilpe linear regression - Combination - CSV - 996 - d:\tmp\mybox-data-6.5.9\data\DiabetesPrediction_standardized_en.csv

Data	Filter	Options	About data analysis					
			Data...	Dependen...	Independent variable	Adjus...▼	Coeffici...	Coeff +
19	disease pr...	[sex, BMI(body mass index), ...]	0.50767...	0.51771...	[-0.14]			
18	disease pr...	[sex, BMI(body mass index), ...]	0.50764...	0.51657...	[-0.14]			
10	disease pr...	[age, sex, BMI(body mass index), ...]	0.50656...	0.51774...	[-0.00]			
9	disease pr...	[age, sex, BMI(body mass index), ...]	0.50650...	0.51657...	[-0.00]			
26	disease pr...	[BMI(body mass index), ...]	0.49165...	0.499727	[0.348]			
27	disease pr...	[BMI(body mass index), ...]	0.49099...	0.50023...	[0.344]			
16	disease pr...	[sex, BMI(body mass index), ...]	0.48634...	0.49332...	[-0.14]			
17	disease pr...	[sex, BMI(body mass index), ...]	0.48585...	0.49401...	[-0.14]			
7	disease pr...	[age, sex, BMI(body mass index), ...]	0.48525...	0.49342...	[0.010]			
8	disease pr...	[age, sex, BMI(body mass index), ...]	0.48477...	0.49411...	[0.010]			
24	disease pr...	[BMI(body mass index), ...]	0.471218	0.47721...	[0.376]			
25	disease pr...	[BMI(body mass index), ...]	0.47023...	0.47744...	[0.376]			
34	disease pr...	[BP(average blood pressure), ...]	0.41056...	0.41991...	[0.238]			
33	disease pr...	[BP(average blood pressure), ...]	0.40788...	0.41594...	[0.256]			
15	disease pr...	[sex, BMI(body mass index), ...]	0.39582...	0.40267...	[-0.04]			
13	disease pr...	[sex, BMI(body mass index), ...]	0.39568...	0.39979...	[-0.06]			
14	disease pr...	[sex, BMI(body mass index), ...]	0.39519...	0.40068...	[-0.06]			
23	disease pr...	[BMI(body mass index), ...]	0.39507...	0.40056...	[0.490]			
4	disease pr...	[age, sex, BMI(body mass index), ...]	0.39477...	0.40026...	[0.023]			
6	disease pr...	[age, sex, BMI(body mass index), ...]	0.39463...	0.40287...	[0.015]			
5	disease pr...	[age, sex, BMI(body mass index), ...]	0.39407...	0.40094...	[0.017]			
21	disease pr...	[BMI(body mass index), ...]	0.39324...	0.39599...	[0.488]			
		..						

Double click selected item to view

2.13.8 Frequency Distributions

1. Select data row. Data filter can be set.
2. Select one column to count frequency.
3. Option whether case-insensitive

4. Set decimal scale.
5. Results can be written as external data, including: new csv/excel/text file, matrix, system clipboard, MyBox clipboard, database table, json, xml, html, or pdf.
6. When handle rows for current page or selected ones, results can be written in current page: insert, append, or replace data in defined location of the table.

The screenshot shows a data table on the left and a configuration dialog on the right.

Data Table:

town	town_Count	town_Count percentage
Beverly	2	3.33
Danvers	4	6.67
Hamilton	1	1.67
Lynn	22	36.67
Lynnfield	2	3.33
Marblehead	3	5
Middleton	1	1.67
Nahant	1	1.67
Peabody	9	15
Salem	7	11.67
Sargus	4	6.67
Swampsc...	2	3.33
Topsfield	1	1.67
Wenham	1	1.67

Frequency Distributions Dialog:

- Column:** town
- Case-insensitive:** checked
- Decimal scale:** 2
- Target:**
 - External:** CSV (selected), XLS, TXT, HTML, JSON, XML, Star, Database, Cache
 - Name:** BostonHousingPrices_en.csv
- In table:** options for Insert above, Append below, Replace

2.13.9 Values Percentage

1. Select data row. Data filter can be set.
2. Select columns to be calculated.
3. Select objects: According to columns/rows/all.
4. Set decimal scale.

5. To negative values: skip, zero, or absolute value.
6. To non-numeric, skip or count as zero.
7. Select columns to be copied
8. Results can be written as external data, including: new csv/excel/text file, matrix, system clipboard, MyBox clipboard, database table, json, xml, html, or pdf.
9. When handle rows for current page or selected ones, results can be written in current page: insert, append, or replace data in defined location of the table.

.6-DataTools-zh.odt - OpenOffice Writer

编辑(E) 视图(V) 插入(I) 格式(O) 表格(A) 工具(T) 窗口(W) 帮助(H)

MyBox Edit Excel File : Excel - Food consumption of China *

Window Document Image Data File Media Network Settings Recent Accessed Development Help

Source row number item year 2020 year 2020_Percent... year 2019 year 2019_Percent...

Source row number	item	year 2020	year 2020_Percent...	year 2019	year 2019_Percent...
Row1	food consumption per capita(kil...	141.2	36.2	130.1	35.12
Row2	cooking oil consumption per ca...	10.4	2.67	9.5	2.56
Row3	vegetables				
Row4	meat consu				
Row5	poultry con				
Row6	aquatic pro				
Row7	eggs consu				
Row8	milk consu				
Row9	fruits consu				
Row10	sugar consu				
Column-Summation					

Value percentage - CSV - 64 - d:\tmp\mybox-data-6.5.9\data\ChinaFoods_en_en.csv

Percentage

Data Row filter Options

Columns to be calculated (No selection means all)

item year 2020 year 2019 year 2018 year 2017 year 2016
 year 2015 year 2014 year 2013

According to Columns Rows All

Decimal scale 2

To negative Skip Count as zero Absolute value

To invalid number Skip Count as zero

Columns to be copied

item year 2020 year 2019 year 2018 year 2017 year 2016
 year 2015 year 2014 year 2013

2.14 Data Charts

Hover or click button “Function” to select functions under menu item “Charts”.

The screenshot shows the MyBox Data Tools interface. A CSV file named "DiabetesPrediction_standardized_en.csv" is being edited. The menu bar includes "Window", "Document", "Image", "Data", "File", "Media", "Network", "Settings", "Recent Accessed", "Development", and "Help". The toolbar features icons for various operations like copy, paste, and search. A context menu is open over the 15th row of a table, with the "Charts" option selected. The "Charts" submenu includes "XY chart", "Pie chart", "Box-and-whisker chart", "Self comparison bars chart" (which is highlighted), "Comparison bars chart", "XYZ chart", and "Location distribution". Other options in the menu include "Data", "Modify", "Trim", "Calculation", "Group Charts", and "Examples". A checked checkbox for "Pop when mouse hovering" is also visible.

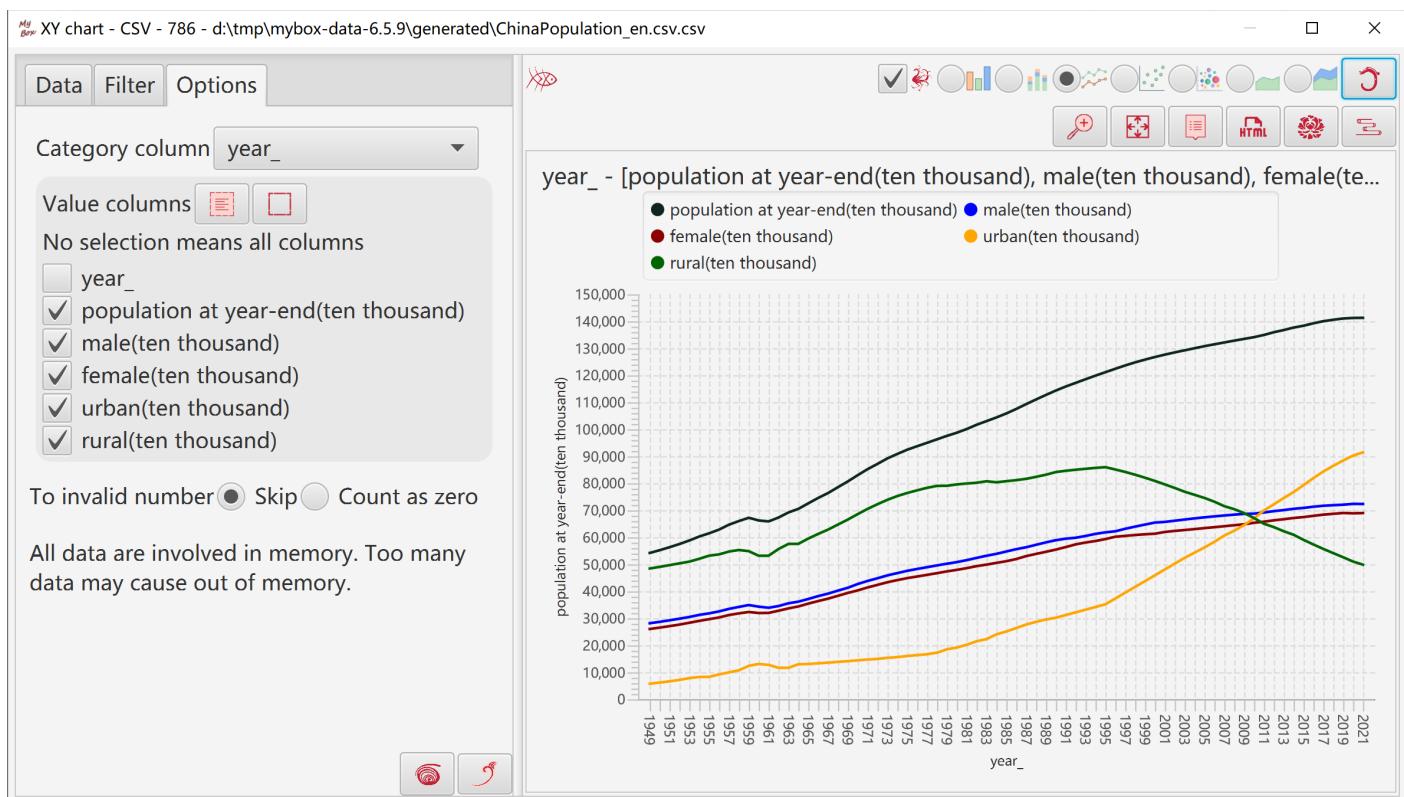
Table r...	Data...	age	sex	BMI(body ...	Bl
1	1	0.038075906	0	0.0616962...	0
2	2	-0.001882...	0	-0.051474...	0
3	3	0.08529890...	0	0.0444512...	0
4	4	-0.089062...	0	0.0444512...	0
5	5	0.00538306...	0	0.0444512...	0
6	6	-0.092695...	0	0.0444512...	0
7	7	-0.045472...	0	0.0444512...	0
8	8	0.06350367...	0	0.0444512...	0
9	9	0.04170844...	0	0.0444512...	0
10	10	-0.070900...	0	0.0444512...	0
11	11	-0.096328...	0	0.0444512...	0
12	12	0.027178291	0	0.0175059...	0
13	13	0.016280676	0	-0.028840...	0
14	14	0.00538306	0	-0.001894...	0
15	15	0.045340983	0	-0.025606...	0

2.14.1 XY Chart

2.14.1.1 Data

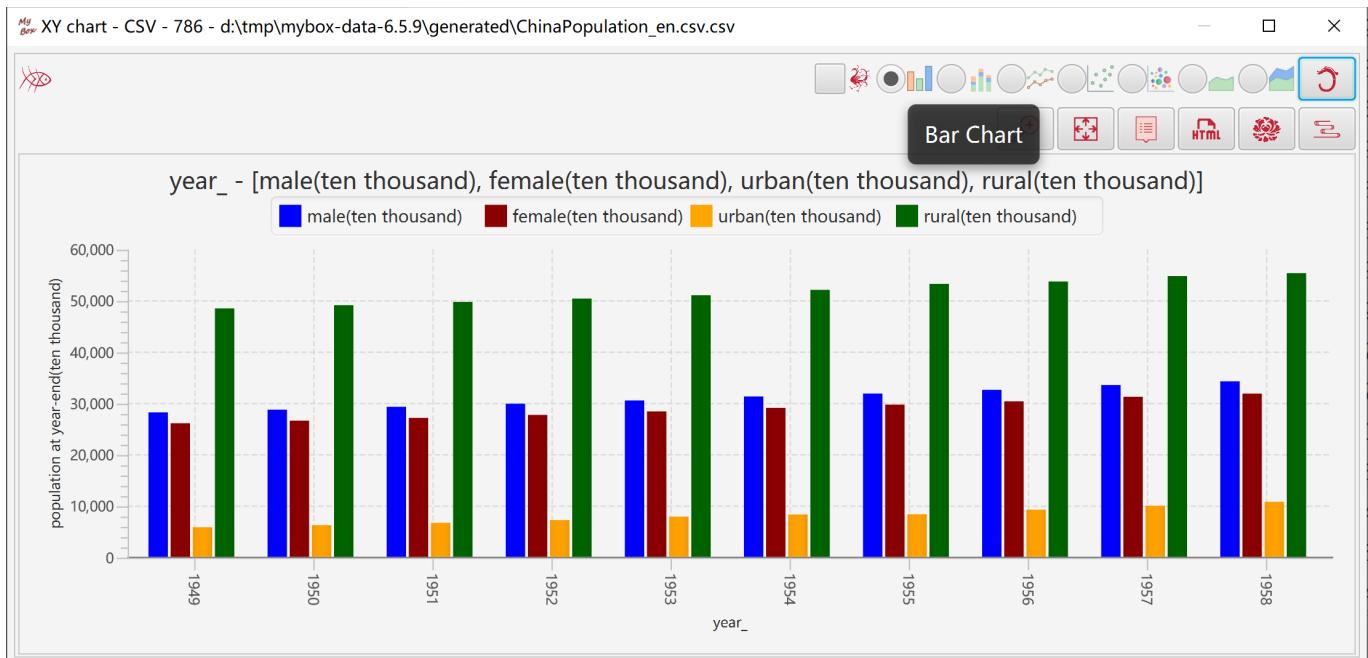
1. Select data rows. Data filter can be set.
2. Select type of XY Chart.

3. Select one column as "Category Axis", to define data names.
4. Select several columns as "Value Axis". Different value series are shown in different colors or shapes.
5. By default, "Category Axis" is the horizontal axis and "Number Axis" is the vertical axis.
6. To invalid data:, skip or count as zero.
7. When display all data rows(all pages), need concern memory limitation.
8. Click button “Menu” to set parameters of chart.
9. Click button “Pop” to display current chart as image in popped window.
10. Click button “Data” to display data of the chart in data table.
11. Click button “Html” to display data of the chart in html page.



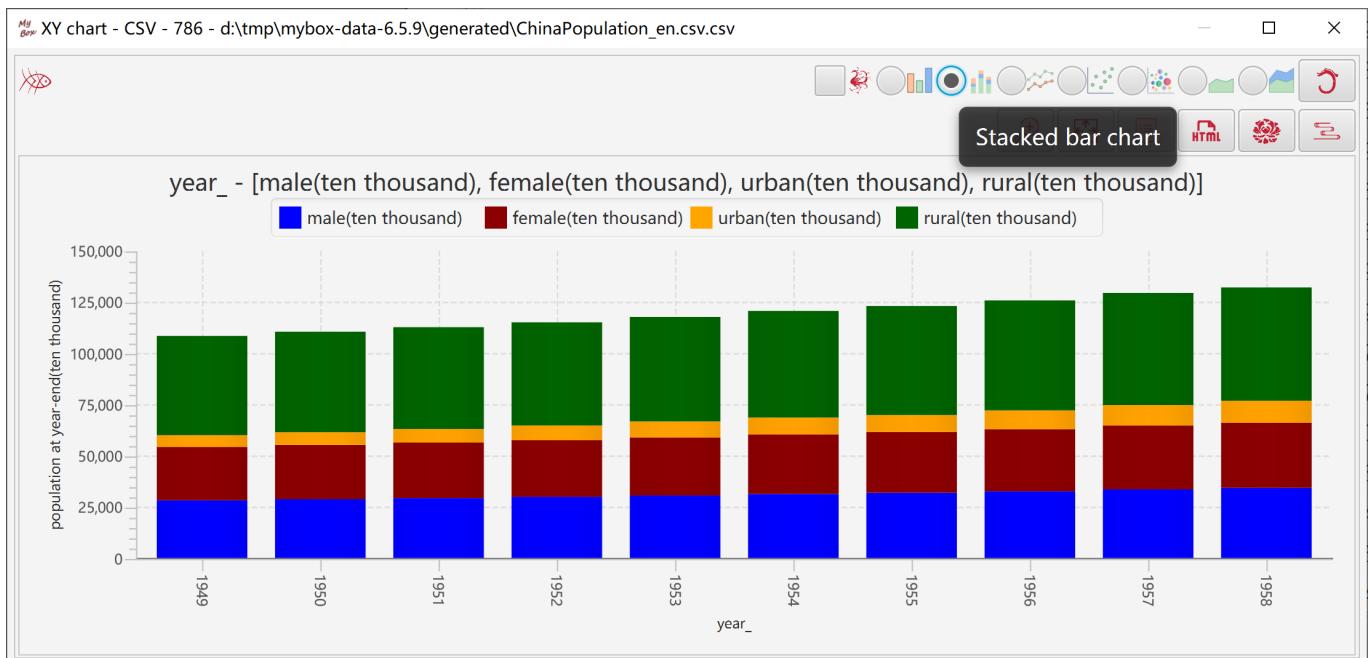
2.14.1.2 Bar Chart

1. Represents data size with bars' heights.
2. "Category" column is always counted as strings.



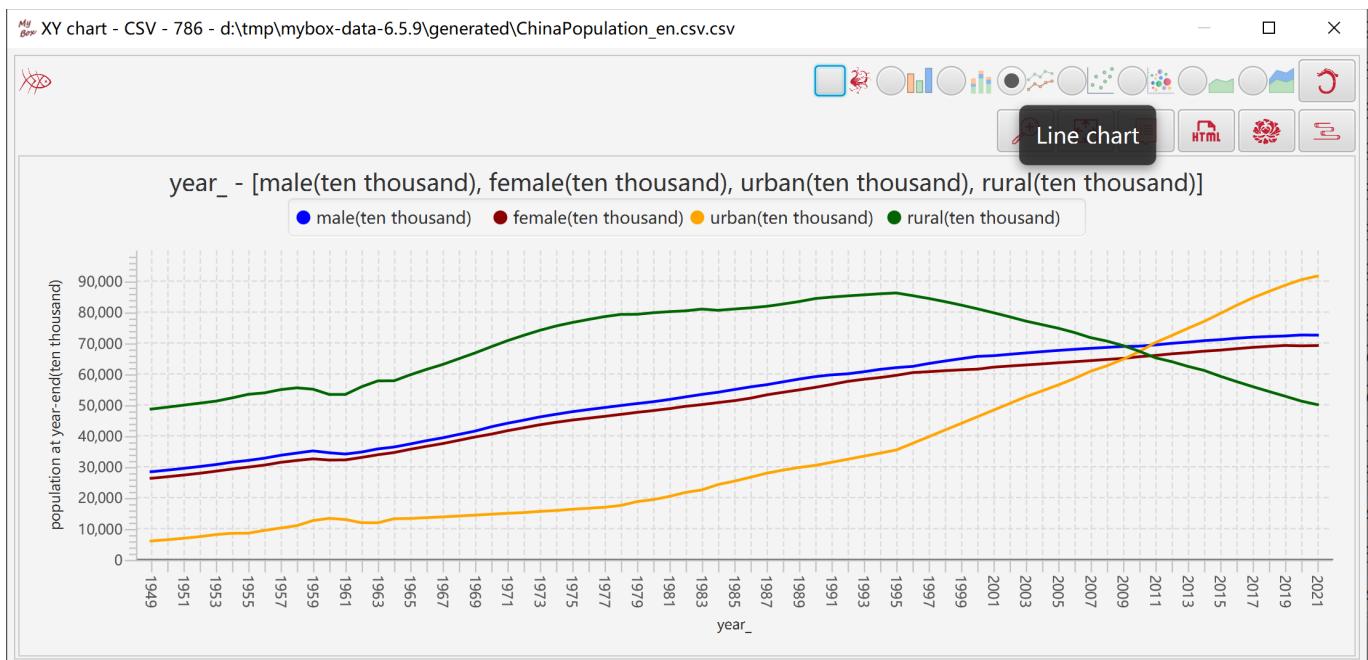
2.14.1.3 Stacked Bar Chart

1. Represents data size with bars' heights.
2. "Category" column is always counted as strings.



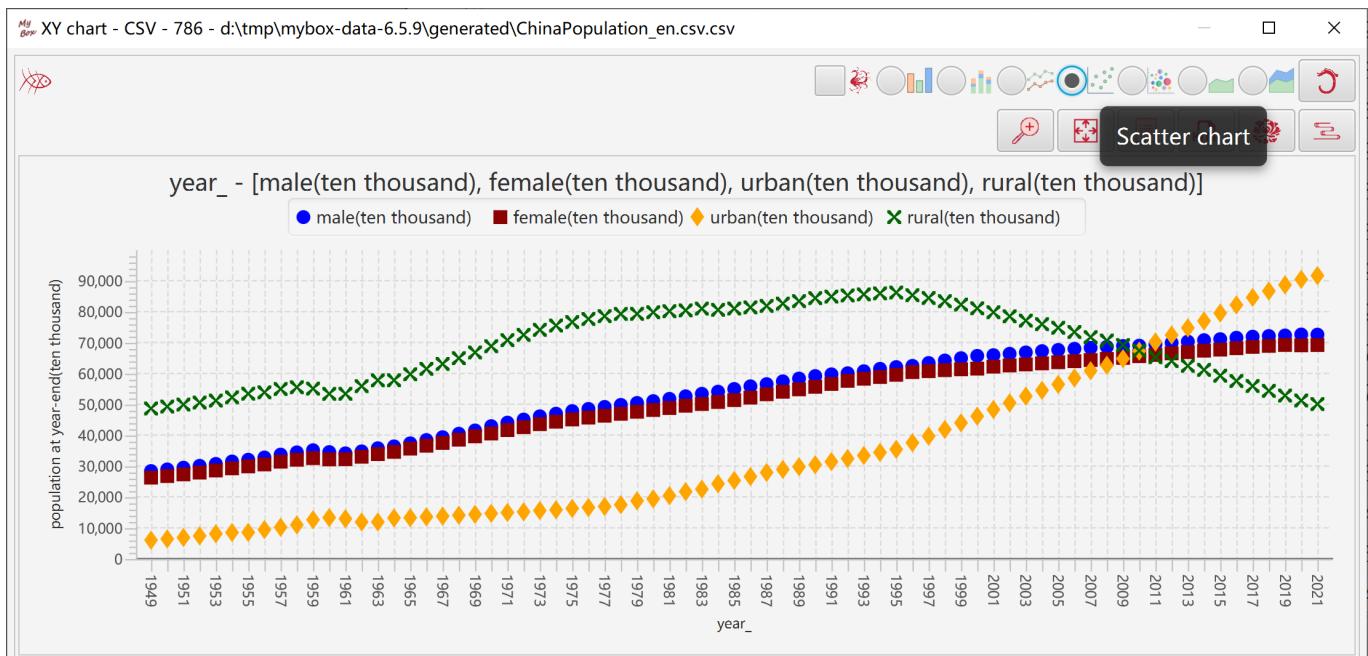
2.14.1.4 Line Chart

1. Represents data trend with lines connecting points.
2. "Category" column can be counted as strings or numbers.



2.14.1.5 Scatter Chart

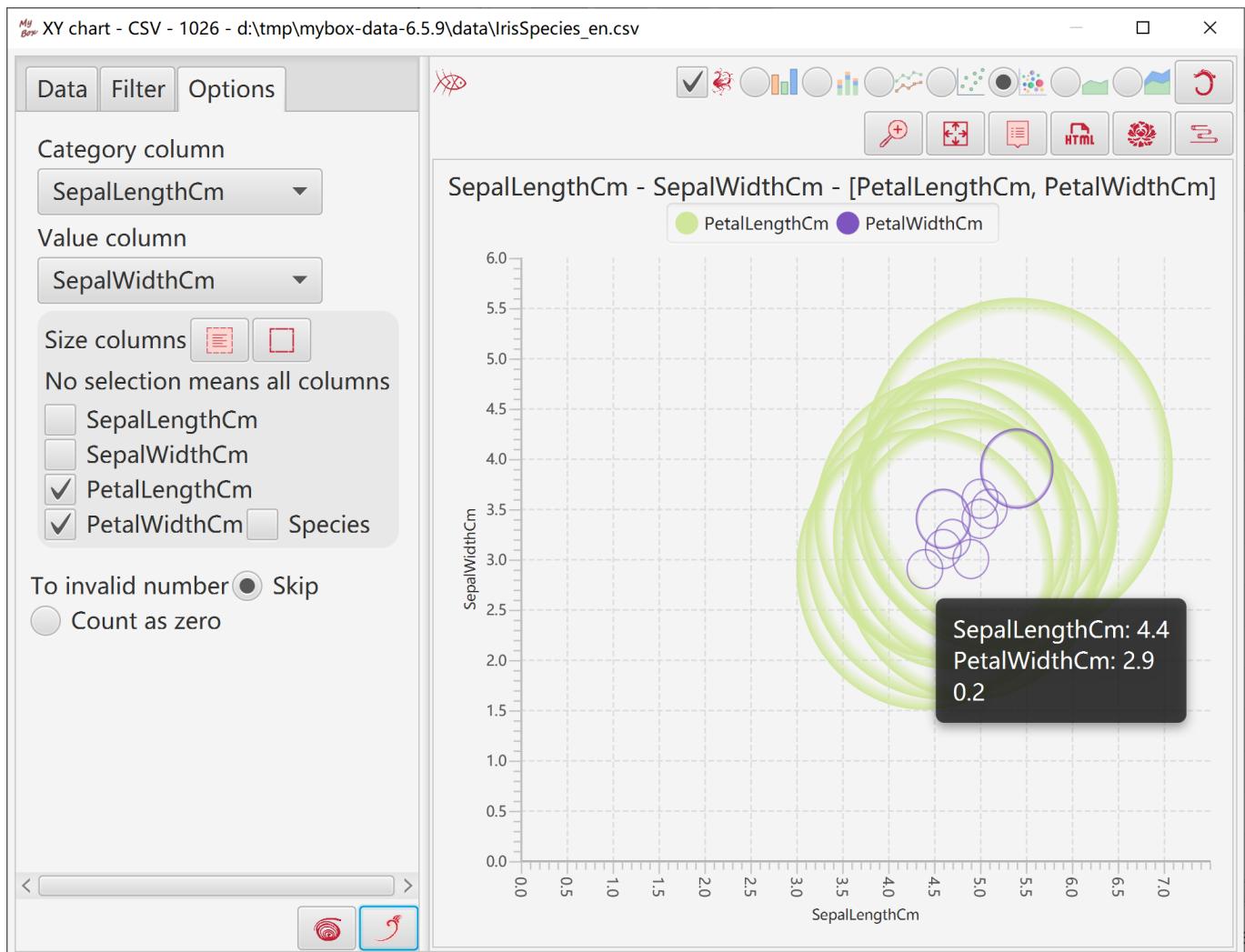
1. Represents data distribution with symbols.
2. "Category" column can be counted as strings or numbers.



2.14.1.6 Bubble Chart

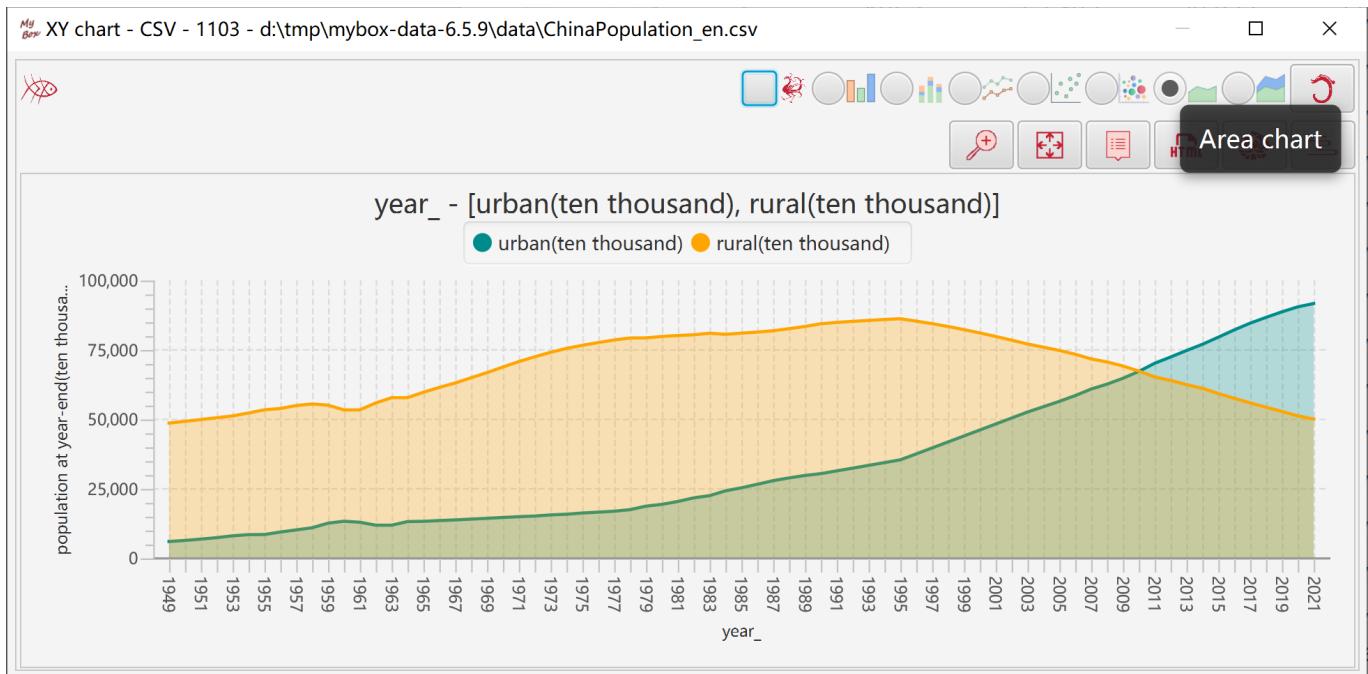
1. Represents data size with circles of different radius:
2. "Category Column" and "Value Column" define coordinates of data.

3. Select several columns as "Size Columns" to defines data size.
4. All columns should be numbers. Size columns should be non-negative.



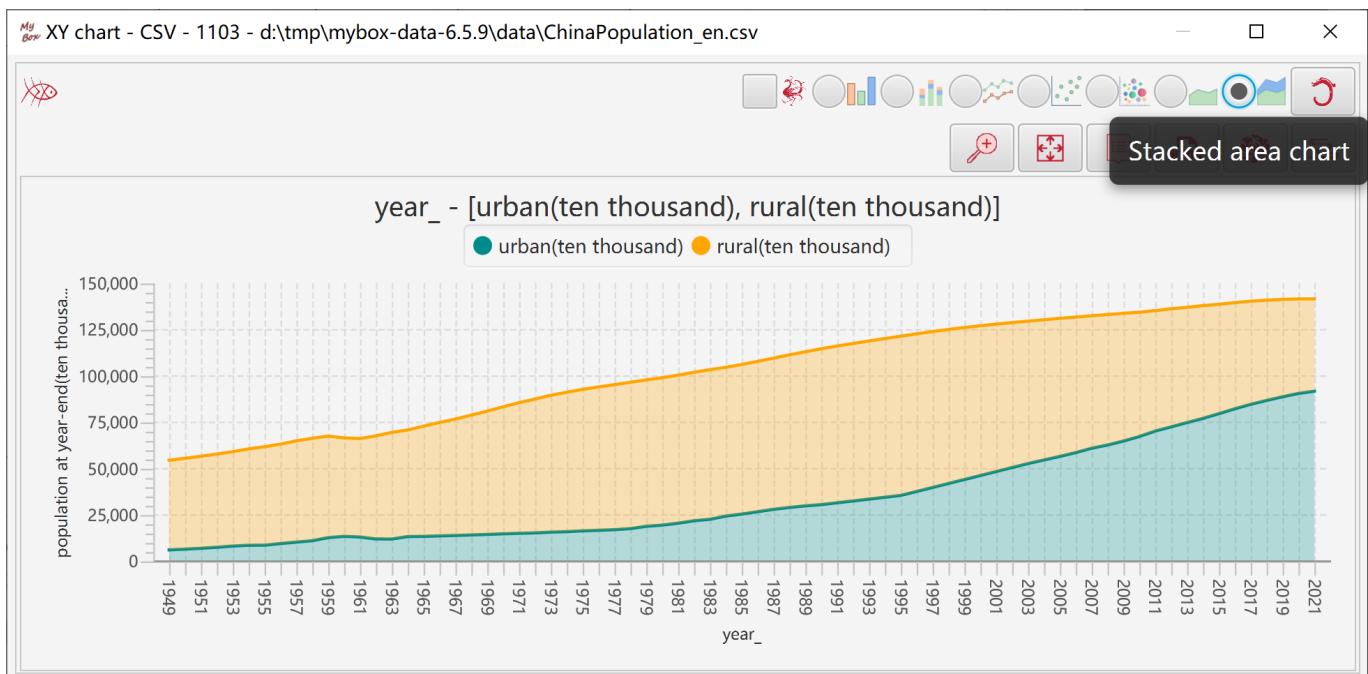
2.14.1.7 Area Chart

1. Represents data size with area size.
2. "Category" column is always counted as strings.



2.14.1.8 Stacked Area Chart

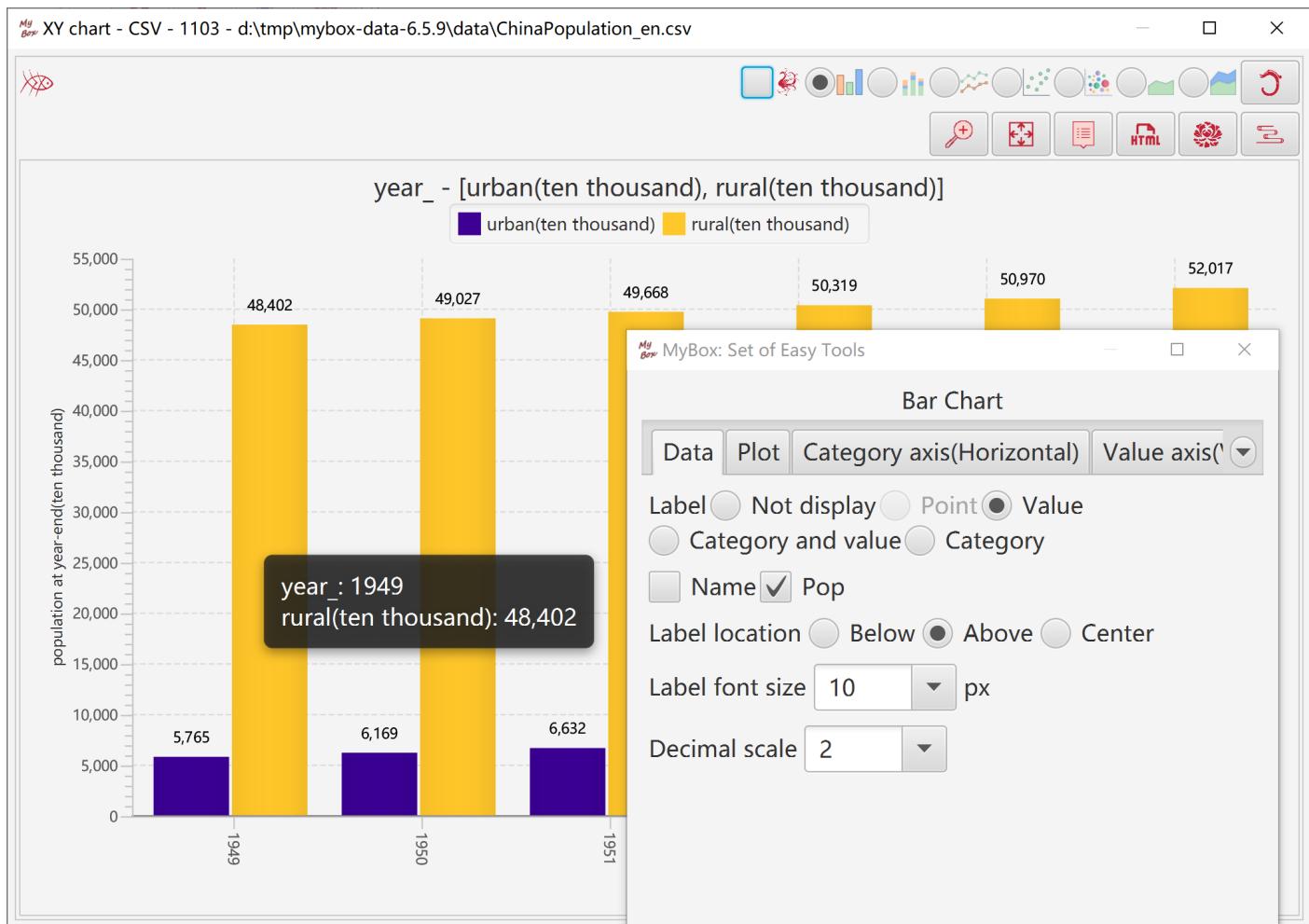
1. Represents data size with area size.
2. "Category" column is always counted as strings.



2.14.1.9 Parameters of Data in Chart

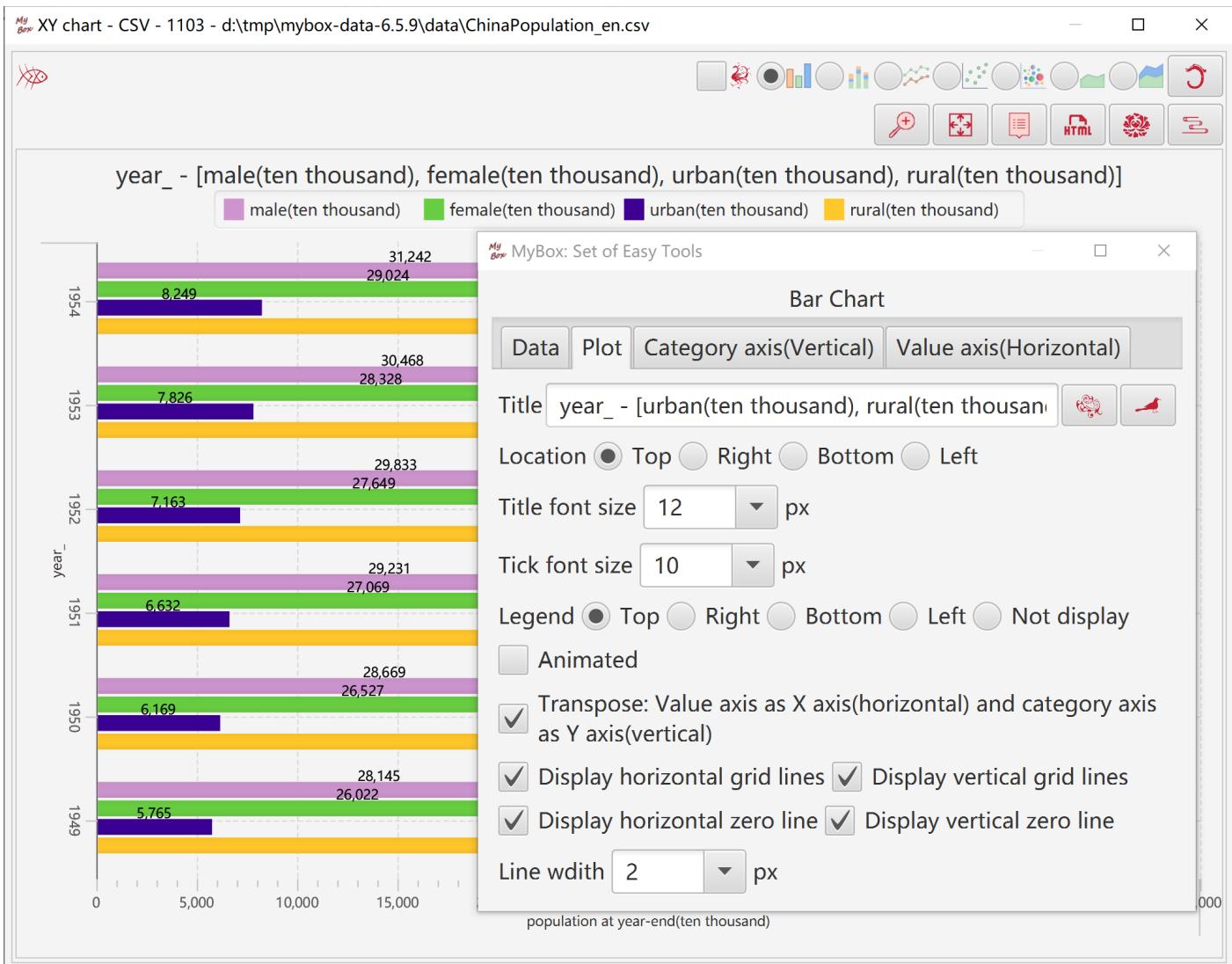
1. Options about label: not display, point, value, category, etc.
2. Location of labels.

3. Font size of labels.
4. Decimal scale.



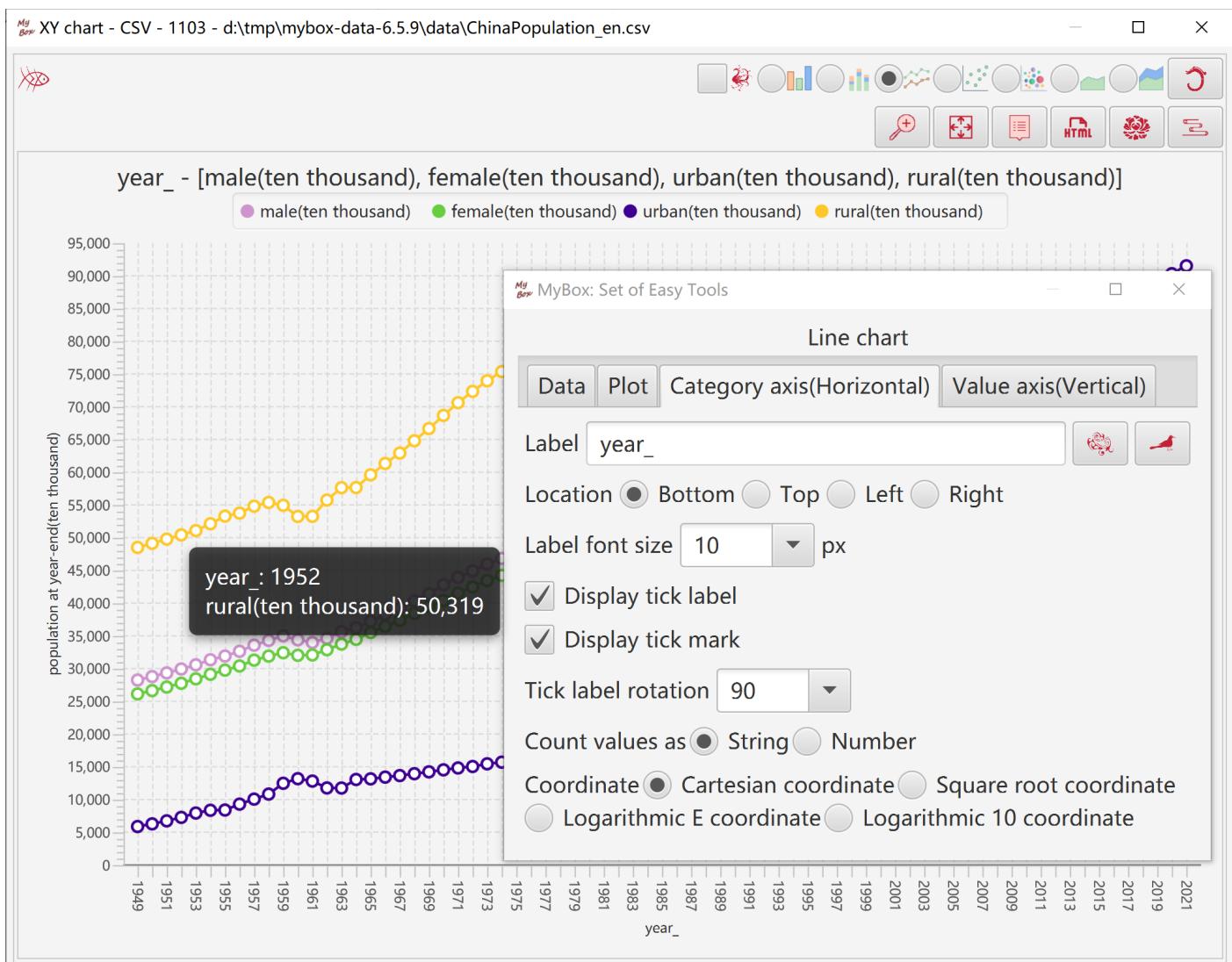
2.14.1.10 Layout

Set parameters of plot: title, font size, location of legend, zero-lines, grid lines, line width, etc.



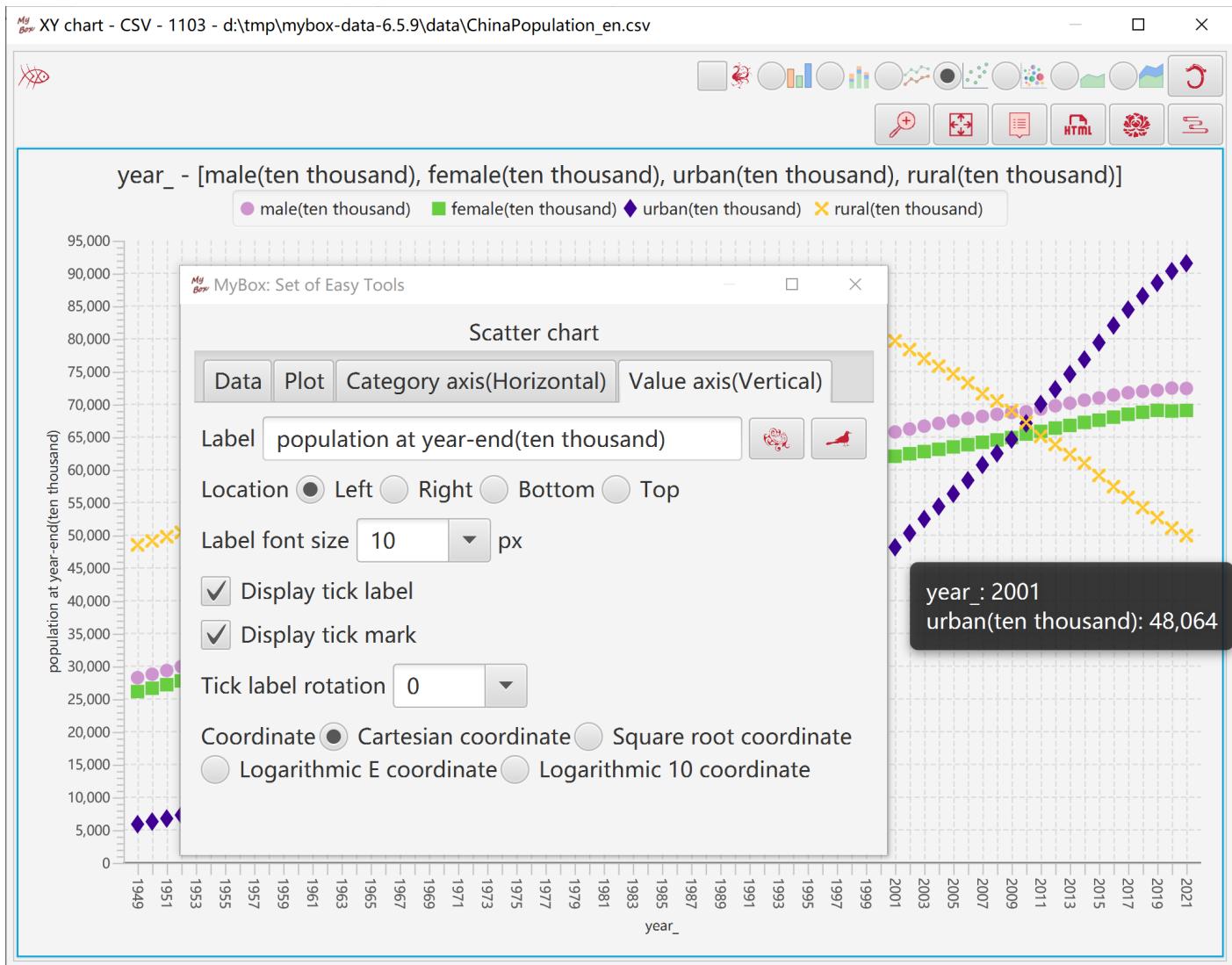
2.14.1.11 Category Axis

Set parameters of category axis: label, font size, location, tick, count values as string or number, coordinate, etc.



2.14.1.12 Number Axis

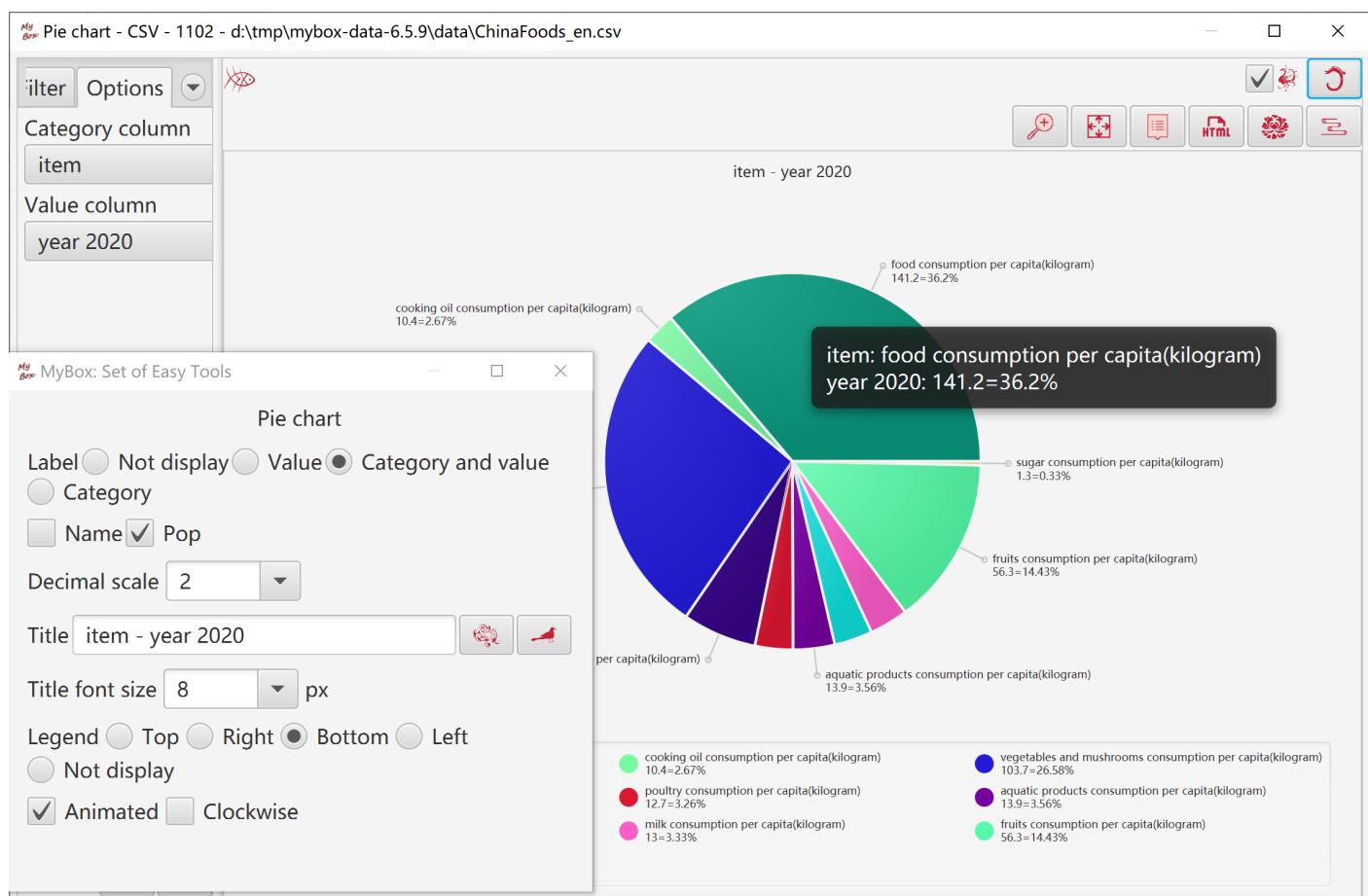
Set parameters of number axis: label, font, location, tick, coordinate, etc.



2.14.2 Pie Chart

Data numbers are represented as percentages with a circle divided into segments.

1. Select data rows. Data filter can be set.
2. Select one column as "Category Axis", to define data names..
3. Select another column as "Number Axis". Values should be non-negative.
4. When display all data rows(all pages), need concern memory limitation
5. Click button “Menu” to set parameters of chart.
6. Click button “Pop” to display current chart as image in popped window.
7. Click button “Data” to display data of the chart in data table.
8. Click button “Html” to display data of the chart in html page.



2.14.3 Box-and-whisker Chart

2.14.3.1 Way of Calculation

Box-and-whisker chart represents data distribution:

1. Sort data according to column/row/all in ascending order.
2. Following items can show aggregation and discreteness of data:

Minimum $Q_0 = \text{in } 0\%(\text{start}) \text{ of the data list}$

Lower quartile $Q_1 = \text{in } 25\% \text{ of the data list}$

Median $Q_2 = \text{in } 50\%(\text{middle}) \text{ of the data list}$

Upper quartile $Q_3 = \text{in } 75\% \text{ of the data list}$

Maximum $Q_4 = \text{in } 100\%(\text{end}) \text{ of the data list}$

3. Following values can be used to mark outliers of data:

Lower extreme outlier line $E_1 = Q_1 - 3 * (Q_3 - Q_1)$

Lower mild outlier line $E_2 = Q_1 - 1.5 * (Q_3 - Q_1)$

Upper mild outlier line $E_3 = Q_3 + 1.5 * (Q_3 - Q_1)$

Upper extreme outlier line $E_4 = Q_3 + 3 * (Q_3 - Q_1)$

4. Following values can be referred for discreteness:

Mean = average of the data list

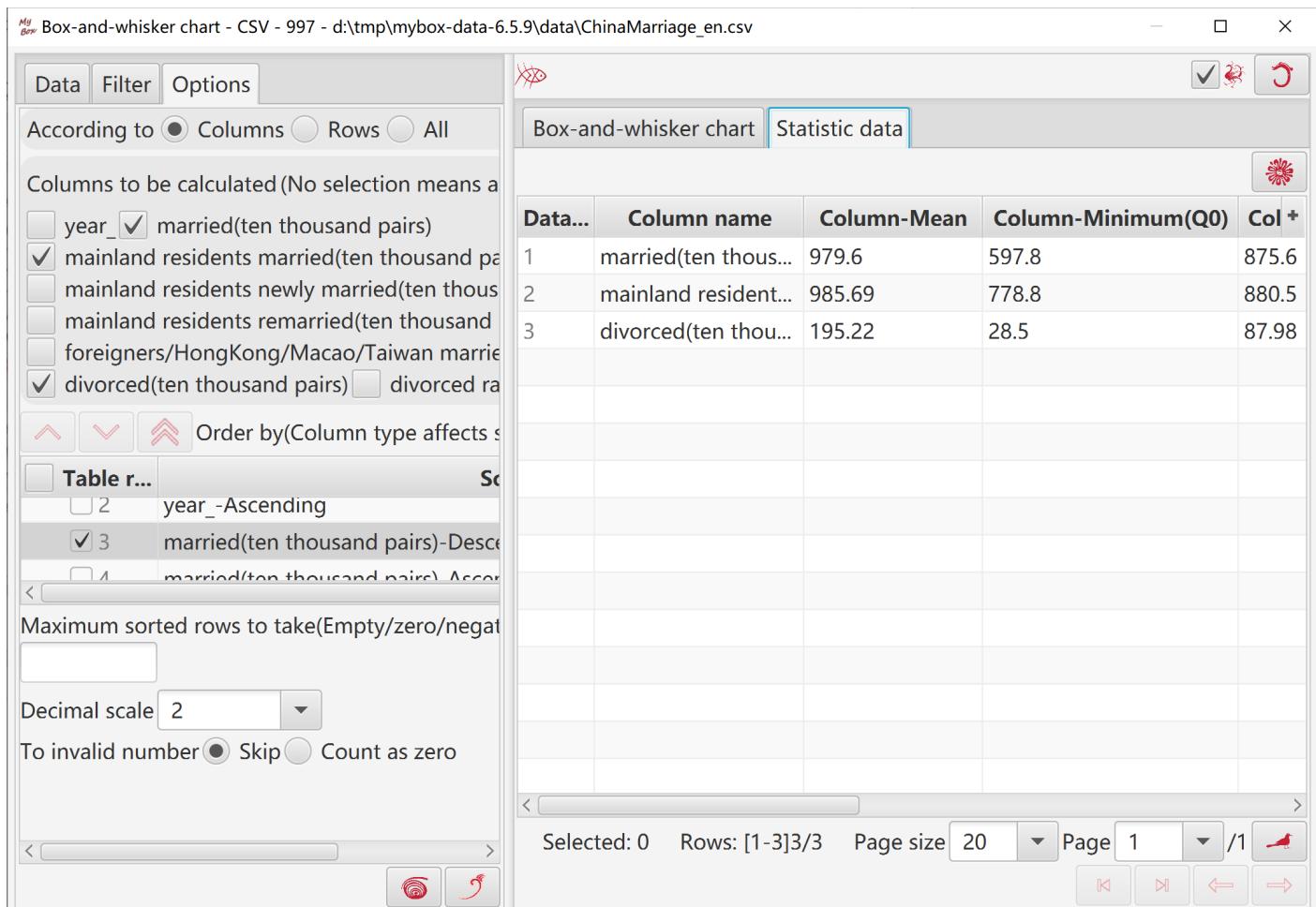
2.14.3.2 Options of Calculation

1. Select data rows. Data filter can be set.
2. Select objects: According to columns, rows(select category column), or all.
3. Select columns to be calculated.

4. Set conditions of sorting, and maximum data number to take.
5. To invalid values: skip, or count as zero.

2.14.3.3 Statistic Data

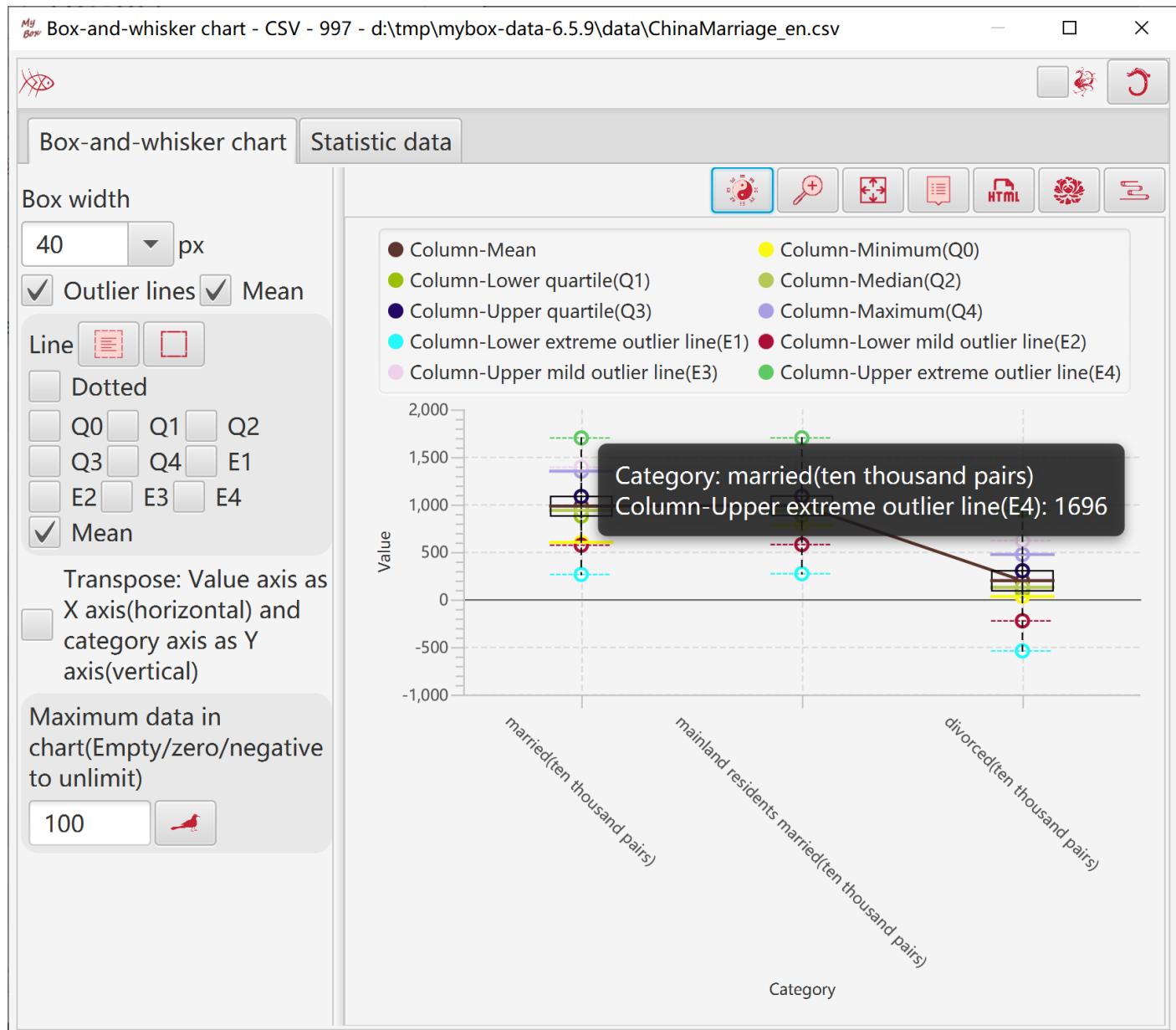
Generated statistic data can be viewed and edited.



2.14.3.4 Options of Chart

1. Box width
2. Whether display outliers lines or mean
3. Whether display connection lines of values, and whether dotted lines

4. Whether transpose
5. Maximum data number in chart
6. Random colors
7. Click button “Menu” to set parameters of chart.
8. Click button “Pop” to display current chart as image in popped window.
9. Click button “Data” to display data of the chart in data table.
10. Click button “Html” to display data of the chart in html page.



2.14.4 Self Comparison Bars Chart

2.14.4.1 Way of Calculation

Self comparison bars show difference between data and reference values. Bars are calculated by following rulers:

1. If value is zero, no bar
2. When compare as absolute values:

maximum_value = maximum_absolute_value_of_column/row/all

percentage = absolute_value / maximum_value

width = maximum_width * percentage

color = If value is larger than zero, color_of_column.

If value is less than zero, inverted_color_of_column

3. When compare as range of minimum and maximum:

maximum_value = maximum_value_of_column/row/all

minimum_value = minimum_value_of_column/row/all

percentage = (value - minimum_value) / (maximum_value - minimum_value)

width = maximum_width * percentage

color = color_of_column

2.14.4.2 Options of Calculation

1. Select data rows. Data filter can be set.
2. Select columns to be calculated.
3. Select columns to be copied.
4. Set conditions of sorting, and maximum data number to take.
5. Select how to handle invalid data.
6. Set decimal scale.
7. When display all data rows(all pages), need concern memory limitation.

2.14.4.3 Options of Chart

1. Select objects to compare: columns/rows/all.
2. Select comparison way: absolute values, or range of minimum and maximum.
3. Set maximum width of bars.
4. Set maximum data number in chart.
5. Select whether display row numbers, values, percentages, categories, calculated values.
6. Edit data in chart.
7. Edit html of chart.

MyBox User Guide – Data Tools v6.7

Self comparison bars chart - CSV - 944 - d:\tmp\mybox-data-6.5.9\data\ChinaPopulation_en.csv

Data Filter Options

Columns to be calculated (No selection means all)

year_ population at year-end(ten thousand) male(ten thousand) female(ten thousand) urban(ten thousand) rural(ten thousand)

Columns to be copied

year_ population at year-end(ten thousand) male(ten thousand) female(ten thousand) urban(ten thousand) rural(ten thousand)

Order by (Column type affects sorting results)

Table r...	Sort
year_ Descending	
2 year_-Ascending	
3 population at year-end(ten thousand)	
4 population at year-end(ten thousand)	

Maximum sorted rows to take (Empty/zero/negative)

Decimal scale 2

To invalid number Skip Count as zero

All data are involved in memory. Too many data may out of memory.

Compare

Columns Rows All

According to

Absolute value Range of min and max

Maximum width 80

Maximum data in chart (Empty/zero/negative to unlimit)

100

Row number Value Percentage Others Calculated values

CSV - 944 - d:\tmp\mybox-data-6.5.9\data\ChinaPopulation_en.csv - Self comparison bars chart Column comparison

male(ten thousand)	female(ten thousand)	year_	male(ten thousand)	female(ten thousand)
Max absolute: 72357.0	Max absolute: 68969.0			
99%	99%	2021	72311	68949
100%	99%	2020	72357	68855
99%	100%	2019	72039	68969
99%	99%	2018	71864	68677
99%	99%	2017	71650	68361
98%	98%	2016	71307	67925
97%	97%	2015	70857	67469
97%	97%	2014	70522	67124
96%	96%	2013	70063	66663

2.14.5 Comparison Bars Chart

2.14.5.1 Way of Calculation

Comparison bars show difference between two series of data. Bars are calculated by following rulers:

1. If value is zero, no bar
2. When compare as absolute values:

maximum_value = maximum_absolute_value_of_value_columns

percentage = absolute_value / maximum_value

width = maximum_width * percentage

color = If value is larger than zero, color_of_column.

If value is less than zero, inverted_color_of_column

3. When compare as range of minimum and maximum:

maximum_value = maximum_value_of_value_columns

minimum_value = minimum_value_of_value_columns

percentage = (value - minimum_value) / (maximum_value - minimum_value)

width = maximum_width * percentage

color = color_of_column

2.14.5.2 Options of Calculation

1. Select data rows. Data filter can be set.
2. Select one column as category column(unnecessary).
3. Select two columns to be compared.
4. Select columns to be copied.
5. Set conditions of sorting, and maximum data number to take.
6. Set decimal scale.
7. When display all data rows(all pages), need concern memory limitation.

2.14.5.3 Options of Chart

1. Select comparison way: absolute values, or range of minimum and maximum.
2. Set maximum width of bars.
3. Set maximum data number in chart.
4. Select whether display row numbers, values, percentages, categories, calculated values.
5. Edit data in chart.
6. Edit html of chart.

Comparison bars chart - CSV - 944 - d:\tmp\mybox-data-6.5.9\data\ChinaPopulation_en.csv

Data **Filter** **Options**

Category column **year_**

Value column **male(ten thousand)**

Value column **female(ten thousa...**

Columns to be copied

year_ population at year-end(ter
 male(ten thousand) female(ten t
 urban(ten thousand) rural(ten th

Order by(Column type affects sorting res

Table r...

1 year_-Descending
 2 year_-Ascending
 3 population at year-end(te
 4 population at year-end/te

Maximum sorted rows to take(Empty/zero)

Decimal scale **2**

All data are involved in memory. Too ma
out of memory.

According to
 Absolute value
 Range of min and max

Maximum width **80**

Maximum data in chart(Empty/zero/negative to unlimit)
100

Row number Value Percentage Category
Others Calculated values

**CSV - 944 - d:\tmp\mybox-data-
6.5.9\data\ChinaPopulation_en.csv -
Comparison bars chart**

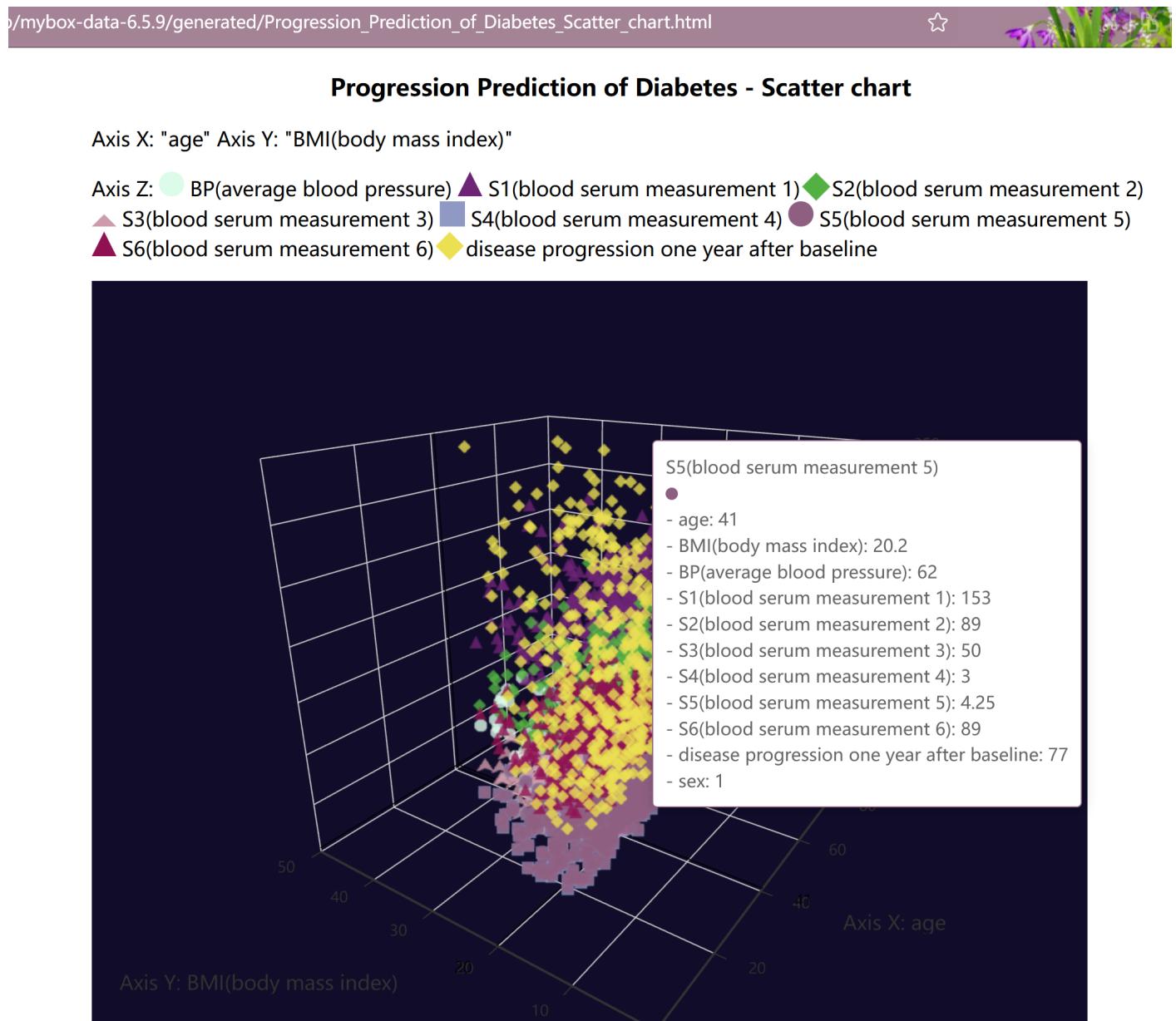
Maximum: 72357.0 Minimum: 26022.0

male(ten thousand)	year_	female(ten thousand)	urban(ten thousand)	rural(ten thousand)
99%	2021	92%	91425	49835
100%	2020	92%	90220	50992
99%	2019	92%	88426	52582
98%	2018	92%	86433	54108
98%	2017	91%	84343	55668
97%	2016	90%	81924	57308
96%	2015	89%	79302	59024
96%	2014	88%	76738	60908
95%	2013	87%	74502	62224
94%	2012	86%	72175	63747

H2

2.14.6 XYZ Chart

1. Base on [echarts-gl](#), 3D charts can be displayed in html with technic of WebGL.
2. Webview of JavaFX does not support WebGL, so generated html will be displayed by system web browser.
3. Please make sure your system web browser supports WebGL and does not limit local JavaScript files.



2.14.6.1 3-D Scatter

1. Select data rows. Data filter can be set.
2. Select data axes:
 - Select one column as axis X. Select as string or number.
 - Select one column as axis Y. Select as string or number.
 - Select several columns as axis Z. Select as string or number.
3. Select other columns to be popped in labels
4. Set parameters of the chart: projection, color, width, height, whether dark, point size.

XYZ chart - CSV - 942 - d:\tmp\mybox-data-6.5.9\data\DiabetesPrediction_en.csv

XYZ chart

Data **Filter** **Options**

Axis X age Count values as strings

Axis Y BMI(body mass in... Count values as strings

Axis Z (No selection means all) Count values as strings

sex BP(average blood pressure) S1(blood serum measurement 1) S2(blood serum measurement 2)
 S3(blood serum measurement 3) S4(blood serum measurement 4) S5(blood serum measurement 5)
 S6(blood serum measurement 6) disease progression one year after baseline

Other values to be popped

age sex BMI(body mass index) BP(average blood pressure) S1(blood serum measurement 1)
 S2(blood serum measurement 2) S3(blood serum measurement 3) S4(blood serum measurement 4)
 S5(blood serum measurement 5) S6(blood serum measurement 6)
 disease progression one year after baseline

Type Scatter chart Surface chart

Projection Perspective Orthographic

Color Column Gradient Random

Width px

Height px

Dark mode

Point size

Notice: A html file will be generated to display the chart.
 Please make sure your system web browser supports WebGL and does not limit local JavaScript files.

2.14.6.2 Surface Chart

1. Select data rows. Data filter can be set.
2. Select data axes:
 - Select one column as axis X. Select as string or number.
 - Select one column as axis Y. Select as string or number.
 - Select one column as axis Z. Select as string or number.
3. Select other columns to be popped in labels.
4. Set parameters of the chart: projection, color, width, height, whether dark, wire frame.

XYZ chart - CSV - 942 - d:\tmp\mybox-data-6.5.9\data\DiabetesPrediction_en.csv

XYZ chart

Data **Filter** **Options**

Axis X BMI(body mass in... Count values as strings

Axis Y S5(blood serum m... Count values as strings

Axis Z disease progressio... Count values as strings

Other values to be popped
 age sex BMI(body mass index) BP(average blood pressure) S1(blood serum measurement 1)
 S2(blood serum measurement 2) S3(blood serum measurement 3) S4(blood serum measurement 4)
 S5(blood serum measurement 5) S6(blood serum measurement 6)
 disease progression one year after baseline

Type Scatter chart Surface chart

Projection Perspective Orthographic

Color Column Gradient Random

Width px

Height px

Dark mode Wire frame

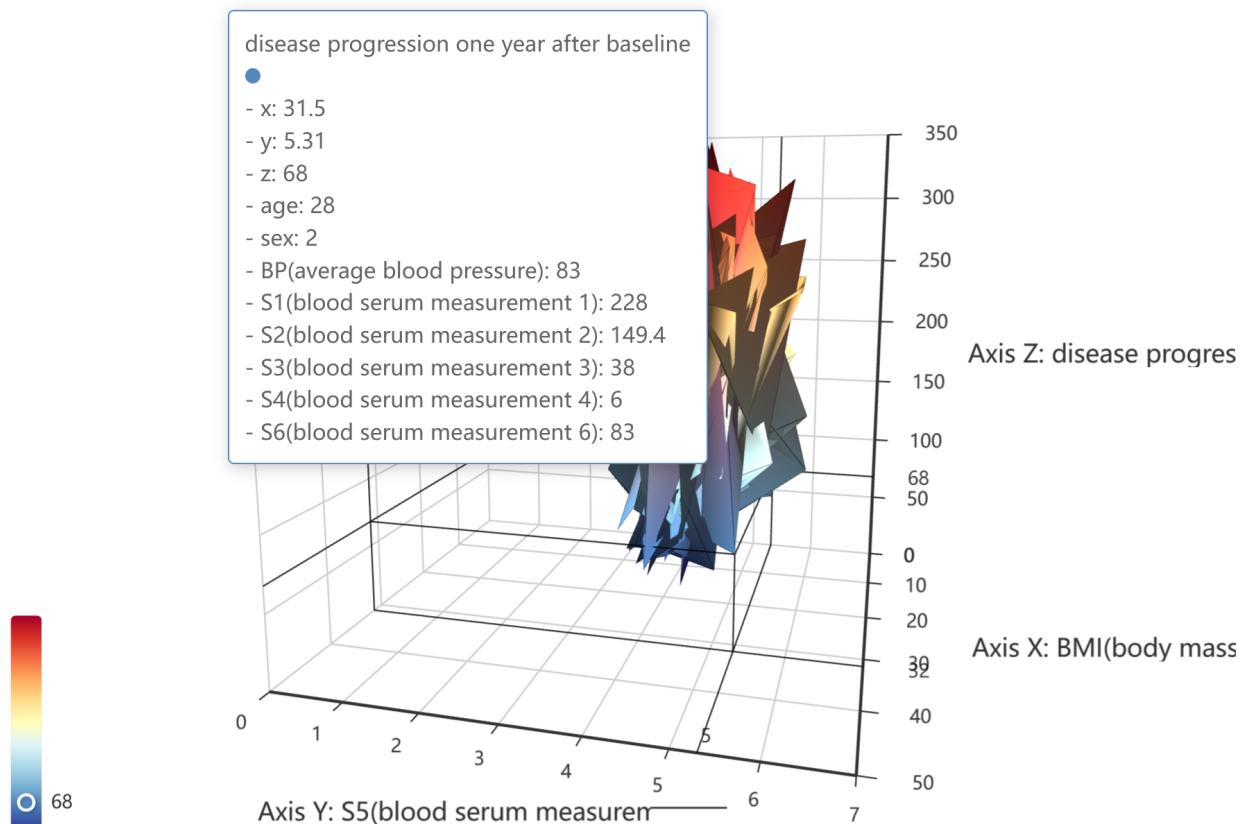
Notice: A html file will be generated to display the chart.
 Please make sure your system web browser supports WebGL and does not limit local JavaScript files.





Progression Prediction of Diabetes - Surface chart

Axis X: "BMI(body mass index)" Axis Y: "S5(blood serum measurement 5)" Axis Z: "disease progression one year after baseline"



2.14.7 Locations Distribution

Display location data in map dynamically.

2.14.7.1 Options of Data

1. Select data rows. Data filter can be set.
2. Select one column for each of following: Label, Longitude, Latitude, and Size.
3. Select other columns to be displayed in popped labels.
4. Select coordinate system of data.
5. Set conditions of sorting, and maximum data number to take.

2.14.7.2 Data in Map

Data referred by map are displayed in table.

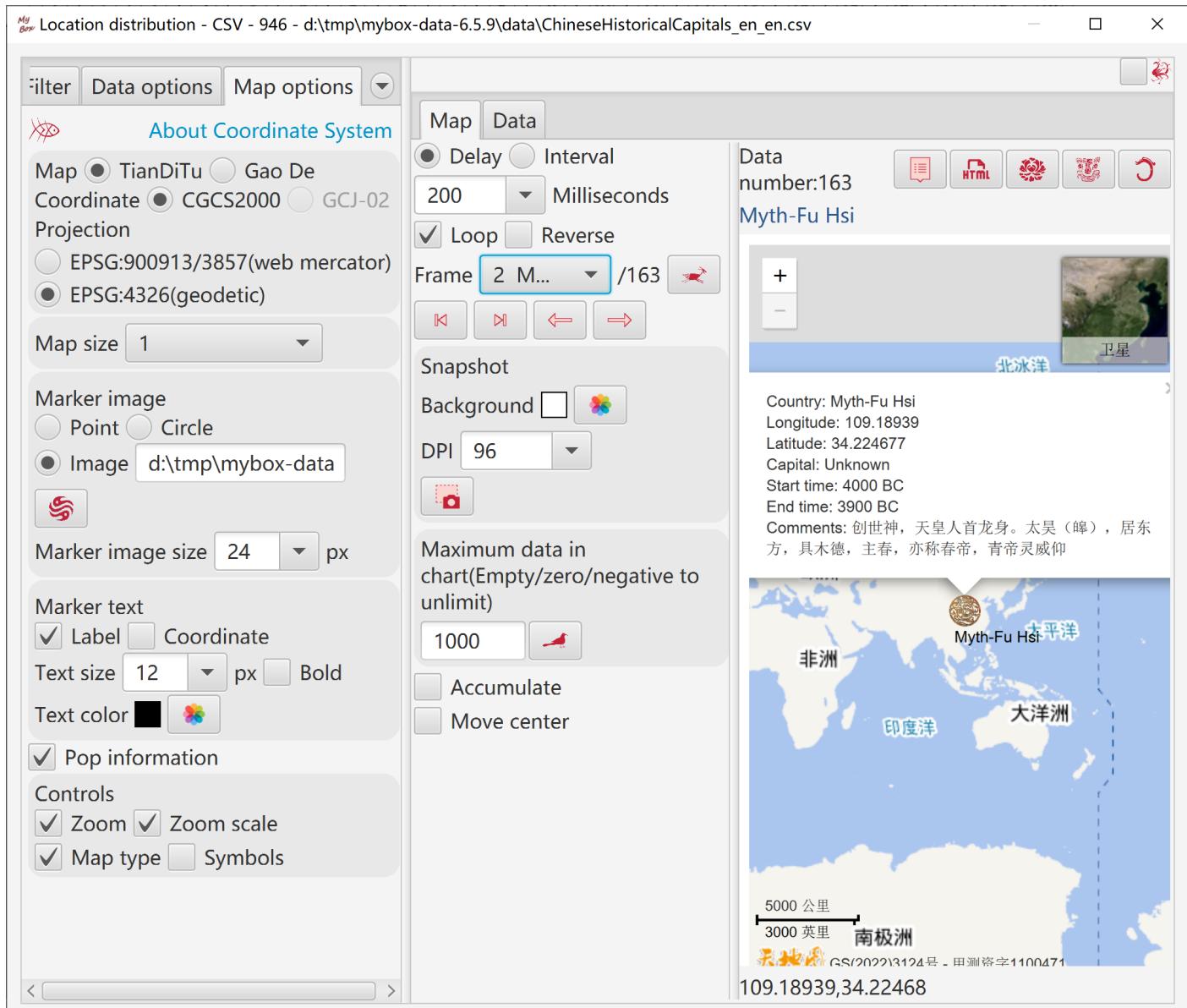
The screenshot shows the 'Location distribution' window in MyBox Data Tools. The left side has a toolbar with 'Data', 'Filter', 'Data options', and 'Map options'. Under 'Data options', settings include 'Label' (Country), 'Longitude' (Longitude), 'Latitude' (Latitude), 'Size' (Not set), and 'Other values to be popped' (Country, Capital, Longitude, Latitude, Start time, End time, Comments). It also includes a 'Coordinate System' section with radio buttons for CGCS2000, GCJ-02, BD-09, and Mapbar, and a 'Sort' section for 'Table r...' (Country-Descending, Country-Ascending). A note says 'All data are involved in memory. Too many data may cause out of memory.' The right side shows a table with 19 rows of data:

Data...	Country	Longitude	Latitude	Comments
1	Myth-Shui Ren	112.55699	34.87973	Unknown
2	Myth-Fu Hsi	109.18939	34.224677	Unknown
3	Myth-Nu Wa	111.107528	35.126412	Unknown
4	Myth-Shen Nun...	113.772655	26.489902	Unknown
5	Myth-Yellow Em...	113.730529	34.395562	Unknown
6	Myth-Shao Hao	118.822224	34.30994	穷桑 (江苏省)
7	Myth-Zhuan Xu	115.071879	34.445656	帝丘 (河南省)
8	Myth-Emperor Ku	115.65637	34.414172	毫 (河南省)
9	Myth-Yao	111.518976	36.088005	唐 (山西省)
10	Myth-Shun	110.329194	34.831075	蒲坂 (山西省)
11	Xia Dynasty	111.220456	35.141363	安邑 (山西省)
12	Xia Dynasty	113.050492	34.453667	阳城 (河南省)
13	Xia Dynasty	113.488478	34.140701	阳翟 (河南省)
14	Xia Dynasty	111.007528	35.026412	安邑 (山西省)
15	Xia Dynasty	112.693124	34.69775	斟鄩 (河南省)
16	Xia Dynasty-Ho...	112.789612	34.90797	鉏/穷石 (河南省)
17	Xia Dynasty	114.28185	34.8138	老丘 (河南省)
18	Xia Dynasty	114.392392	36.097577	西河 (河南省)
19	Shang Dynasty	115.65537	34.454172	亳 (河南省)

At the bottom, there are buttons for 'Selected: 0', 'Rows: [1-20]20/163', 'Page size 20', and a page navigation bar with buttons for first, previous, next, and last pages.

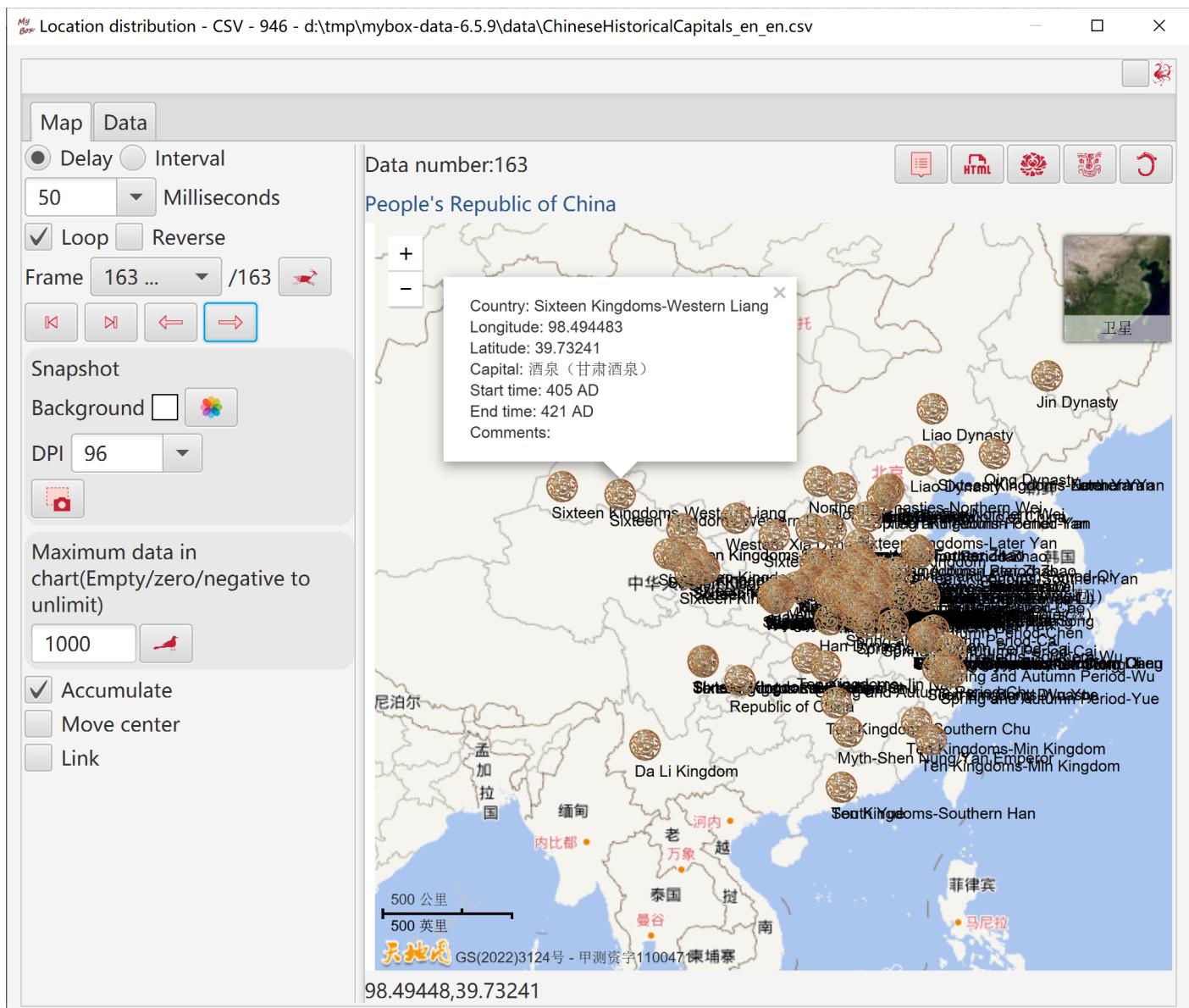
2.14.7.3 Options of Map

1. Map type: TianDi map, or GaoDe map.
2. Select projection and size of map.
3. Set image and text of markers.
4. Select whether pop information.
5. Select controls of map.



2.14.7.4 Options of Playing

1. Select timer:
 - Delay: After display current frame, wait for defined time, and then display next frame.
 - Interval: Keep defined time between current frame and next frame.
2. Select whether: Loop, Reverse, Accumulate, Move Center, Link.
3. Auto-play.
4. Display specific frame.
5. Buttons of navigation.
6. Snapshot:
 - Select background color and dpi.
 - Snap frame by frame, and display them in “Edit Images List”.
7. Set maximum data number in chart.



2.15 Group Charts

Hover or click button “Function” to select functions under menu item “Group Charts”.

The screenshot shows a MyBox Data Tools interface with a CSV file named "ConcreteCompressiveStrength_en.csv" loaded. The file contains data for 15 rows, each with columns for Cement(kg), Blast Furnace Slag(kg), Fly Ash(kg), and other variables. A context menu is open over the data grid, with the "Edit" menu expanded. The "Group Charts" option is highlighted in blue, indicating it is selected. Other options in the "Edit" menu include "Data", "Modify", "Trim", "Calculation", "Charts", "Examples", and "Pop when mouse hovering". The status bar at the bottom shows "Selected: 0 Rows: [1-50]50/1030 Page size 50 Page 1 /21".

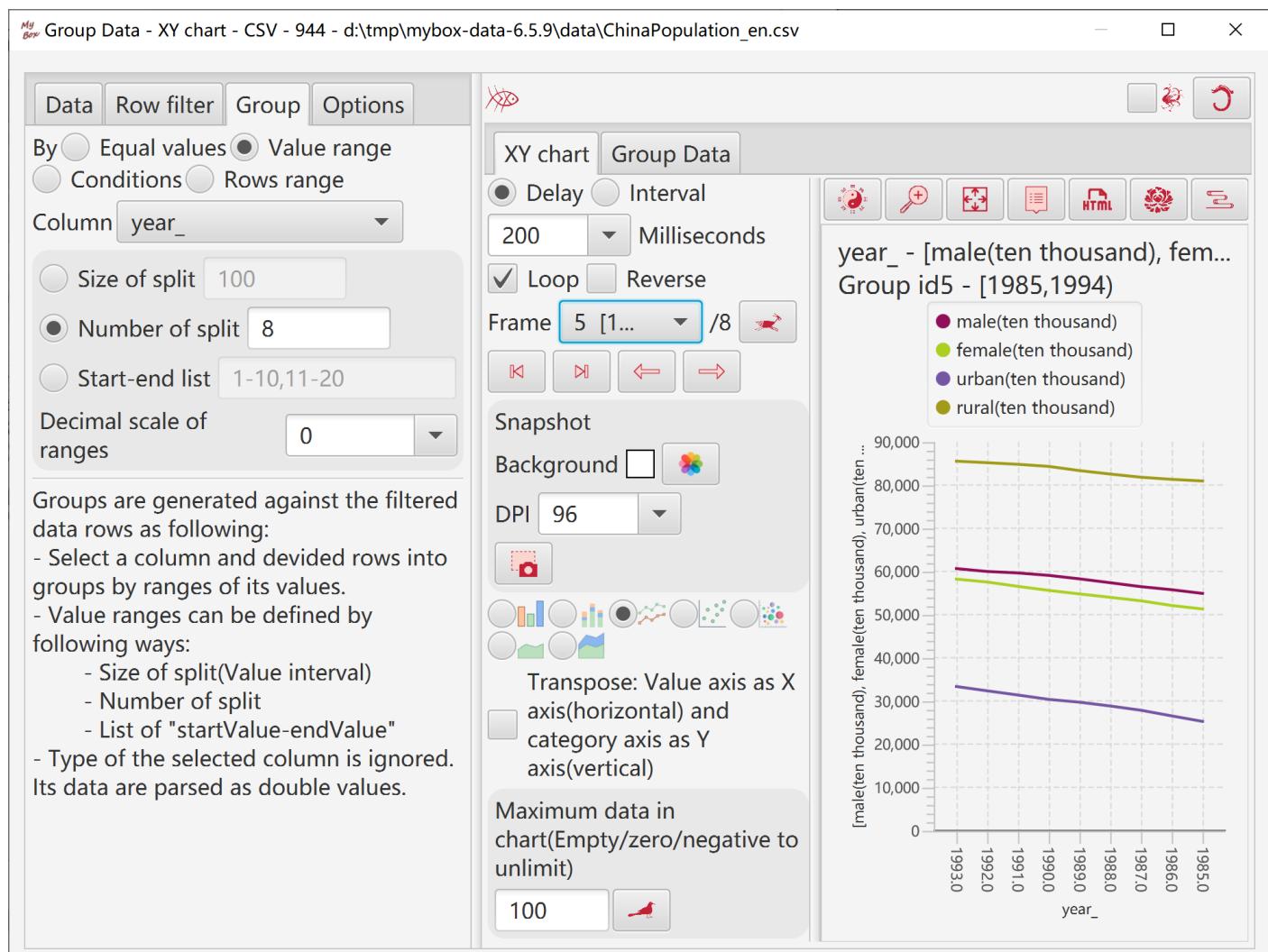
Table r...	Data...	Cement(kg)	Blast Furn...	Fly Ash(kg)	V...
1	1	540	0	0	16
2	2	540	0	0	16
3	3	333	143	0	22
4	4	333	143	0	22
5	5				
6	6				
7	7				
8	8				
9	9				
10	10	475	0	0	228
11	11	199	132	0	192
12	12	199	132	0	192
13	13	428	48	0	228
14	14	190	190	0	228
15	15	304	76	0	228

2.15.1 General Manufacture

After data are grouped, generate one chart for each group, and display them in sequence dynamically.

2.15.1.1 Group Data

1. Select data rows. Data filter can be set.
2. Set conditions of grouping.
3. Set specific options to generate chart.
4. Set conditions of sorting, and maximum data number to pick.
5. Set decimal scale, and how to handle invalid values.



2.15.1.2 Results of Grouping

After calculation, grouped data are displayed in table.

The screenshot shows the 'Group Data - XY chart - CSV - 944' window in MyBox Data Tools. The left panel contains configuration options for grouping:

- Category column:** year_
- Value columns (No selection means all):**
 - year_
 - population at year-end(ten thousand)
 - male(ten thousand)
 - female(ten thousand)
 - urban(ten thousand)
 - rural(ten thousand)
- Order by (Column type affects sorting results):**

Table r...	Sort
<input type="checkbox"/> 1	year_-Descending
<input type="checkbox"/> 2	year_-Ascending
<input checked="" type="checkbox"/> 3	population at year-end(ten thousand)
<input type="checkbox"/> 4	population at year-end(ten thousand)
- Maximum rows in each group (Empty/zero/negative):** (empty input field)
- Decimal scale:** 2
- To invalid number:** Set as blank (radio button selected)
- Row number:** (checkbox)

The right panel displays the grouped data table:

Data...	id	group_id	range_year_	year_	male_...	fer...
1	1	1	[1949,1958)	1957	33469	311
2	2	1	[1949,1958)	1956	32536	302
3	3	1	[1949,1958)	1955	31809	296
4	4	1	[1949,1958)	1954	31242	290
5	5	1	[1949,1958)	1953	30468	283
6	6	1	[1949,1958)	1952	29833	276
7	7	1	[1949,1958)	1951	29231	270
8	8	1	[1949,1958)	1950	28669	265
9	9	1	[1949,1958)	1949	28145	260
10	10	2	[1958,1967)	1966	38189	363
11	11	2	[1958,1967)	1965	37128	354
12	12	2	[1958,1967)	1964	36142	343
13	13	2	[1958,1967)	1963	35533	336
14	14	2	[1958,1967)	1962	34517	327
15	15	2	[1958,1967)	1959	34890	323

Bottom navigation and status bar:

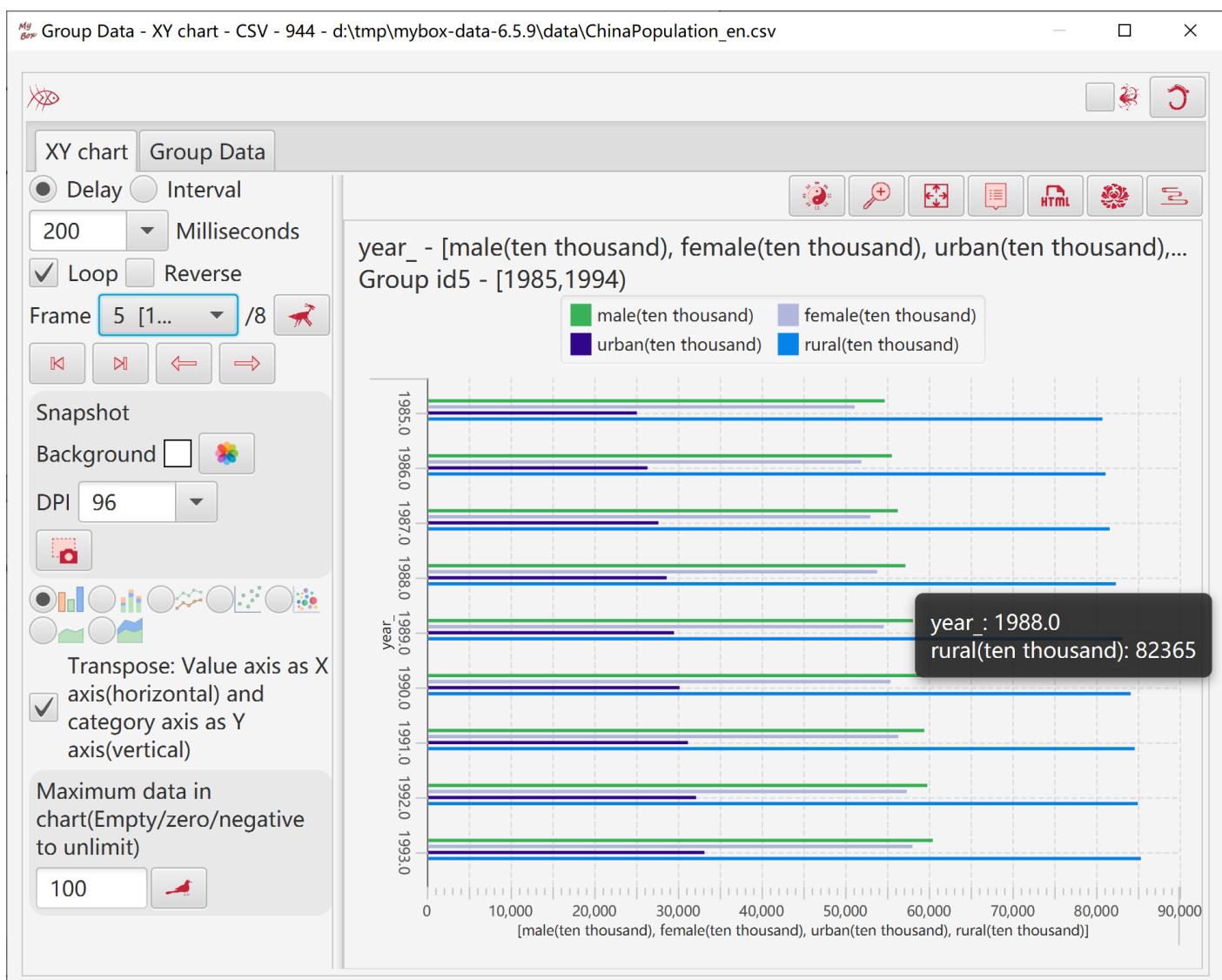
- Selected: 0 Rows: [1-20] 20/73 Page size 20
- Page navigation buttons: 1, /4, back, forward, last, first
- Page: 113 / Total 164

2.15.1.3 Display Options of Chart

1. Specific display options of chart.
2. Set maximum data number displayed in chart.

2.15.1.4 Options of Playing

1. Select timer:
 - Delay: After display current frame, wait for defined time, and then display next frame.
 - Interval: Keep defined time between current frame and next frame.
2. Select whether: Loop, Reverse, Accumulate, Move Center, Link.
3. Auto-play.
4. Display specific frame.
5. Buttons of navigation.
6. Snapshot:
 - Select background color and dpi.
 - Snap frame by frame, and display them in “Edit Images List”.



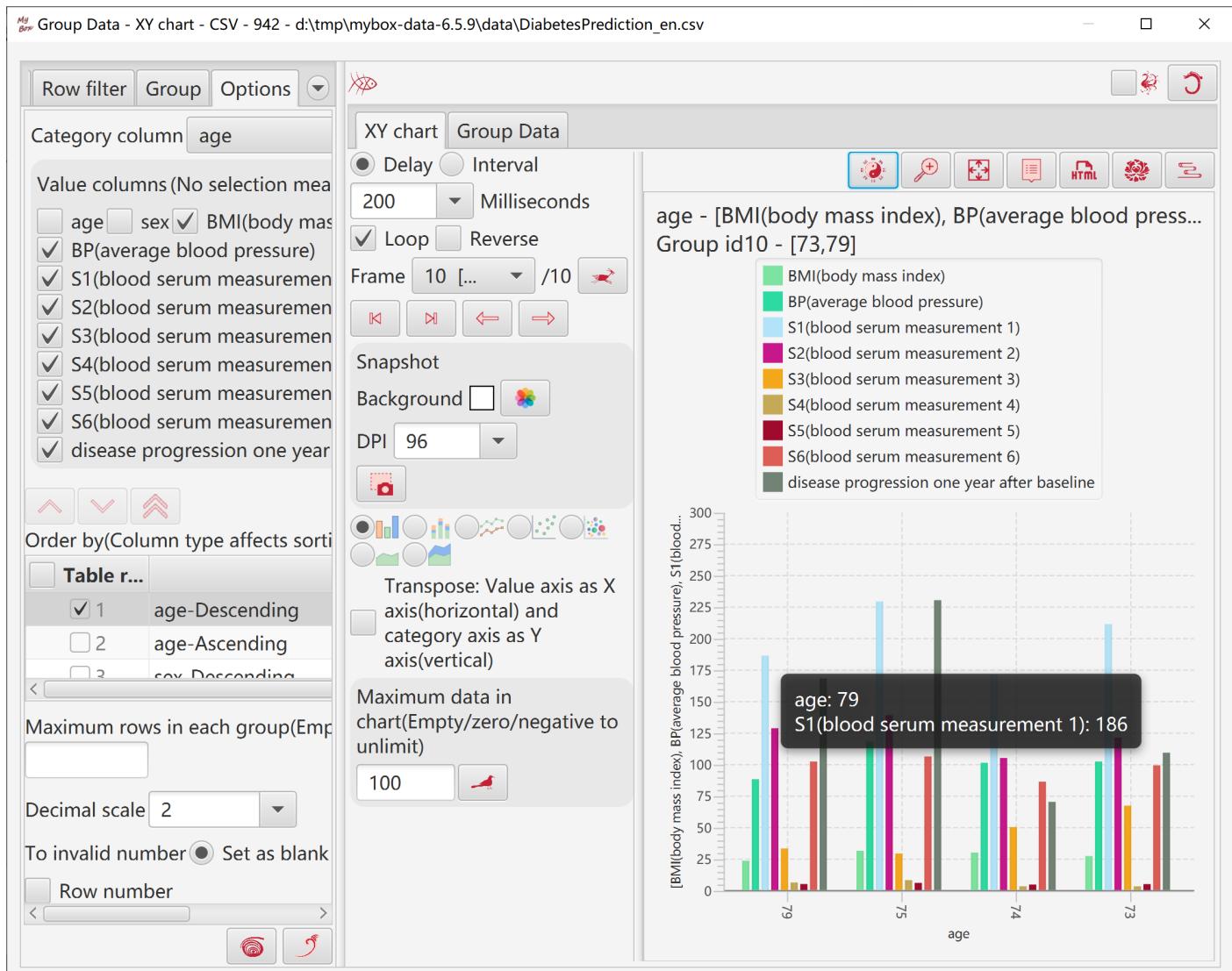
2.15.2 Group Data – XY Chart

2.15.2.1 Specific Options to Generate Chart

1. Select one column as category axis.
2. Select several columns as value axis.

2.15.2.2 Specific Options to Display Chart

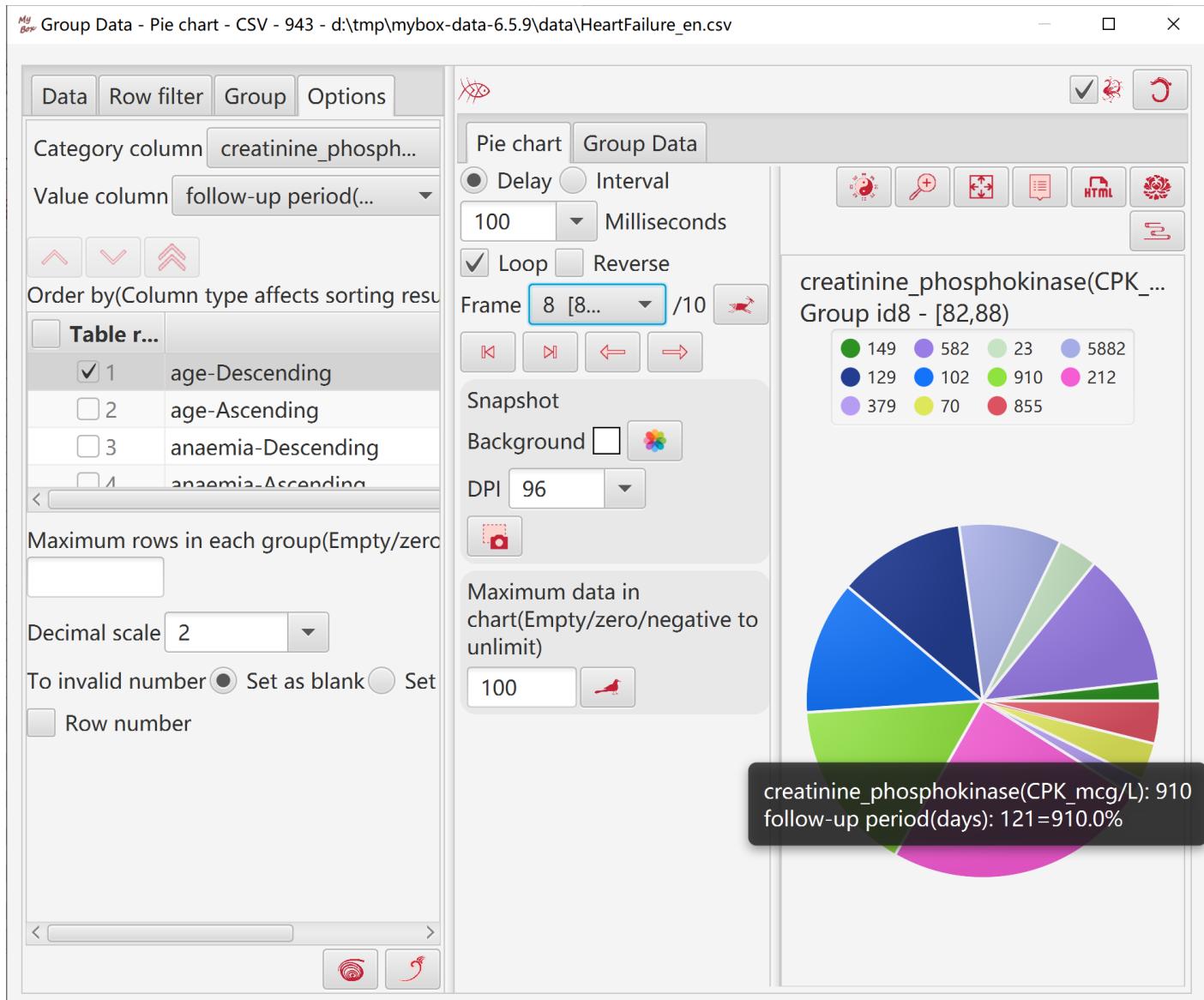
1. Type of XY chart
2. Whether transpose



2.15.3 Group Data – Pie Chart

2.15.3.1 Specific Options to Generate Chart

1. Select one column as category axis.
2. Select one column as value axis.



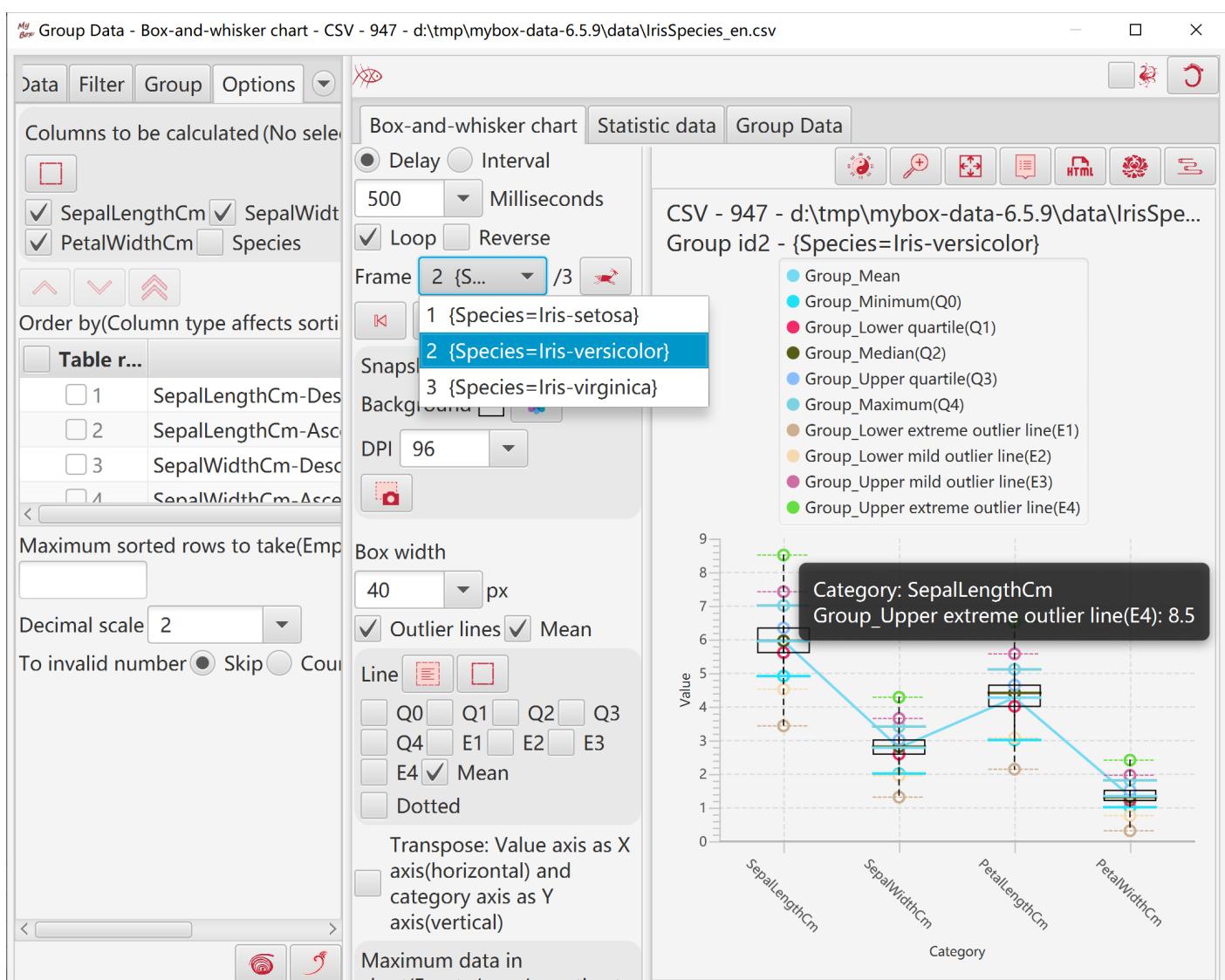
2.15.4 Group Data – Box-and-whisker Chart

2.15.4.1 Specific Options to Generate Chart

1. Select several columns to be calculated.

2.15.4.2 Specific Options to Display Chart

1. Box width
2. Whether display outliers lines or mean
3. Whether display connection lines of values, and whether dotted lines
4. Whether transpose
5. Maximum data number in chart



2.15.5 Group Data – Self Comparison Bars Chart

2.15.5.1 Specific Options to Generate Chart

1. Select several columns to be calculated.
2. Select columns to be copied.

2.15.5.2 Specific Options to Display Chart

1. Select objects to compare: columns/rows/all.
2. Select comparison way: absolute values, or range of minimum and maximum.
3. Set maximum width of bars.
4. Set maximum data number in chart

Group Data - Self comparison bars chart - CSV - 944 - d:\tmp\mybox-data-6.5.9\data\ChinaPopulation_en.csv

Self comparison bars chart

Delay: 500 Milliseconds
 Loop Reverse
Frame: 5 [1... /10

Snapshot

Background:
DPI: 96

Calculated values

CSV - 944 - d:\tmp\mybox-data-6.5.9\data\ChinaPopulation_en.csv - Self comparison bars chart
Group id5 - [1977,1984) Column comparison

urban(ten thousand)	rural(ten thousand)	year_	urban(ten thousand)	rural(ten thousand)
Maximum: 22274.0	Maximum: 80734.0			
Minimum: 16669.0	Minimum: 78305.0			
100%	100%	1983.0	22274	80734
85%	76%	1982.0	21480	80174
62%	65%	1981.0	20171	79901
44%	51%	1980.0	19140	79565
32%	30%	1979.0	18495	79047
1%	29%	1978.0	17245	79014

Columns to be calculated (No selection)

year_ population at year-end(t
 male(ten thousand) female(ten thousand)
 urban(ten thousand) rural(ten thousand)

Columns to be copied

year_ population at year-end(t
 male(ten thousand) female(ten thousand)
 urban(ten thousand) rural(ten thousand)

Order by (Column type affects sorting rule)

Table r...
 1 year_-Descending
 2 year_-Ascending
 3 population at year-end(t
 4 population at year-end(t

Maximum rows in each group (Empty/zero/negative to unlimit)

Decimal scale: 2

To invalid number: Set as blank ?

2.15.6 Group Data – Comparison Bars Chart

2.15.6.1 Specific Options to Generate Chart

1. Select one column as category column(unnecessary).
2. Select two columns to be compared.
3. Select columns to be copied.

2.15.6.2 Specific Options to Display Chart

1. Select comparison way: absolute values, or range of minimum and maximum.
2. Set maximum width of bars.
3. Set maximum data number in chart

Group Data - Comparison bars chart - CSV - 944 - d:\tmp\mybox-data-6.5.9\data\ChinaPopulation_en.csv

Comparison bars chart | Group Data

Delay Interval
500 Milliseconds
Loop Reverse
Frame 5 [1...] /10

Row number Value Percentage Category
Others Calculated values

CSV - 944 - d:\tmp\mybox-data-6.5.9\data\ChinaPopulation_en.csv - Comparison bars chart Group id5 - [1977,1984)

Max absolute: 80734.0

urban(ten thousand)	year_	rural(ten thousand)	urban(ten thousand)	rural(ter thousand)
27%	1983.0	100%	22274	80734
26%	1982.0	99%	21480	80174
24%	1981.0	98%	20171	79901
23%	1980.0	98%	19140	79565
22%	1979.0	97%	18495	79047
21%	1978.0	97%	17245	79014
20%	1977.0	96%	16669	78305

* Links and controls may be unresponsive when html page is editable.

2.16 Input Data

2.16.1 Load Contents in System Clipboard

1. Read contents in System Clipboard.
2. Guess delimiter and parse data.
3. User can select a delimiter to parse data.
4. Select how to parse:
 - CSV parser: Values can contain delimiter and line break. But not support parsing “Blanks” and “Regular Expression”.
 - Text parser: Values can not contain delimiter and line break. Support parsing “Blanks” and “Regular Expression”.
5. Select whether first line defines column names.
6. Based on parsed data, select rows and columns to load. Data filter can be set.

Load content in system clipboard. : CSV - 1032 - d:\tmp\mybox-data-6.5.9\AppData\Temp\2022-11-16_10-...

CSV Texts

First line defines the columns' names

Delimiter: ,

Table r...	Data...	Column...	Column...	Cc +
1	1	age	sex	BMI(b
2	59		2	32.1
3	48		1	21.6
4	72		2	30.5
5	24		1	25.3
6	50		1	23
7	23		1	22.6
8	36		2	22
9	66		2	26.2
10	60		2	32.1
11	29		1	30
12	22		1	18.6

If a value contains delimiter or line break, it should be surrounded by quotes.

Selected: 0 Rows: [1-50]50/51 Page size 50

Page 1 /2 ⏪ ⏩ ⏪ ⏩

2.16.2 Import Examples

Hover or click button “Examples” and item in popped menu.

2.16.2.1 Personal Data

Record private data.

	Year	Population	Area	Other
1	1952	57482	29833	
2	1953	58796	30468	
3	1954	60266	31242	
4	1955	61465	31809	29656 8285
5	1956	62828	32536	30292 9185
6	1957	64653	33469	31184 9949
7	1958	65994	34195	31799 10721
8	1959	67207	34890	32317 12371
9	1960	66207	34283	31924 13073
10	1961	65859	33880	31979 12707
11	1962	67296	31517	32778 11659
12				55636
13				
14				

Selected: 0 Rows: [1-50]50/73 Page size 50 Page 1 /2

2.16.2.2 Statistic Data of China

Data from China National Bureau of Statistics.

The screenshot shows the MyBox Data Tools interface with a CSV file titled "Edit CSV File : CSV - 85 - D:\tmp\mybox-data\mybox-data-6.6.2\data\ChinaMuseums_en.csv". The main window displays a table of data with columns: Year, Type, Number, and another unnamed column. A context menu is open over the table, showing options like "My Data", "Statistic data of China", "Regression data", "Location Data", "Pop when mouse hovering", and "Close(ESC/F6 Or click anywhere outside the object)". The "Statistic data of China" option is currently selected. To the right of the table, a sidebar lists various categories: Population of China, Census of China, Gross domestic product(GDP) of China, Consumer price index(CPI) of China, Food consumption of China, Graduates of China, Museums of China (which is also selected), Health personnel of China, Marriage of China, Sport world champions of China, Crimes filed by China police, Crimes filed by China procuratorate, and China National Bureau of Statistics. The bottom of the interface shows navigation controls for pages, rows, and columns.

	Year	Type	Number	
11	64	2012	5069	71748
12	65	2013	3473	79075
13	66	2014	3658	83970
14	67	2015	3852	89133
15	68	2016	4109	93431
16	69	2017	4721	105079
17	70	2018	4918	107506
18	71	2019	5132	107993
19	72	2020	5452	118913
20	73	2021	3671	

2.16.2.3 Data of Regression

Data can be used to analyse models of regression.

The screenshot shows the MyBox Data Tools interface with a CSV file titled "ConcreteCompressiveStrength_en.csv" open. The interface has a top menu bar with "Window", "Document", "Image", "Data", "File", "Media", "Network", "Settings", "Recent Accessed", "Development", and "Help". Below the menu is a toolbar with various icons. On the left, there's a sidebar with a "Table" icon and a list of data sources: "My Data", "Statistic data of China", "Regression data" (which is selected and highlighted in blue), "Location Data", "Pop when mouse hovering" (with a checked checkbox), and "Close(ESC/F6 Or click anywhere outside the object)". The main area displays a table of concrete compressive strength data. The table has columns for ID, Compressive Strength, Age, and other variables. A dropdown menu from the "Regression data" item lists several datasets: "Income and happiness", "Years experience and salary", "Iris species", "Progression Prediction of Diabetes", "Progression Prediction of Diabetes - standardized", "Heart failure", "Concrete compressive strength" (which is also highlighted in blue), "Dog's radiographs after surgery", "Salaries of baseball players", "South German credit", and "Boston housing prices". At the bottom, there are pagination controls showing "Selected: 0 Rows: [1-50]50/1030 Page size 50 Page 1 /21".

ID	Compressive Strength	Age	Other Variables
4	333	143	0
5	199	132	0
6	266	114	0
7	380	95	0
8	380	95	0
9	266	114	0
10	475	0	0
11	199	132	0
12	199	132	0
13	428	48	0

2.16.3 Location Data

Data can be used to generate “Location Distribution”.

The screenshot shows the MyBox Data Tools interface with a CSV file titled "EuropeanGadwalls_en.csv" open. The file contains data about European Gadwalls, including columns for Longitude, Latitude, and coordinates. A tooltip for the "Location Data" column header is visible, showing options like "Chinese Historical Capitals", "Autumn movement patterns of European Gadwalls", and "Sperm whales Gulf of Mexico". The interface includes a toolbar with various icons for file operations, a menu bar, and a bottom navigation bar with page controls.

		Longitude	Latitude	Coordinate
1				
2	<input checked="" type="checkbox"/> Pop when mouse hovering			
3	<input type="checkbox"/> Close(ESC/F6 Or click anywhere outside the window)			
4	4	2009-08-04 11:00:00	2009-08-04 11:00:00	11.73433 48.21033 WGS_84
5	5	2009-08-04 17:00:00	2009-08-04 17:00:00	11.73317 48.20967 WGS_84
6	6	2009-08-04 18:00:00	2009-08-04 18:00:00	11.733 48.20867 WGS_84
7	7	2009-08-04 23:00:00	2009-08-04 23:00:00	11.733 48.20867 WGS_84
8	8	2009-08-05 09:00:00	2009-08-05 09:00:00	11.74133 48.21067 WGS_84
9	9	2009-08-05 11:00:00	2009-08-05 11:00:00	11.73917 48.21133 WGS_84
10	10	2009-08-05 17:00:00	2009-08-05 17:00:00	11.74067 48.21117 WGS_84
11	11	2009-08-05 18:00:00	2009-08-05 18:00:00	11.74083 48.211 WGS_84
12	12	2009-08-06 11:00:00	2009-08-06 11:00:00	11.7435 48.21117 WGS_84
13	13	2009-08-06 18:00:00	2009-08-06 18:00:00	11.74333 48.21033 WGS_84

2.17 Output Data

Hover or click button “Function” to select functions under menu item “Data”.

The screenshot shows the MyBox Data Tools interface for editing a CSV file named "DiabetesPrediction_en.csv". The window title bar reads "Edit CSV File : CSV - 61 - d:\tmp\mybox-data-6.5.9\data\DiabetesPrediction_en.csv". The menu bar includes "Window", "Document", "Image", "Data", "File", "Media", "Network", "Settings", "Recent Accessed", "Development", and "Help". Below the menu is a toolbar with icons for Save, Recover, Refresh, Open, Create data, Load content in system clipboard, Export, and Convert to database table. A context menu is open over a table, with the "Data" option highlighted. The menu items under "Data" are "Modify", "Trim", "Calculation", "Charts", "Examples", and "Pop when mouse hovering". The main area displays a table with 14 rows and 9 columns, representing the Diabetes dataset. The bottom of the screen shows status information: "Selected: 0", "Rows: 50/442", "Page size: 50", "Page: 1", and navigation buttons for previous/next page and first/last page.

Ta...	Data ...	a						
1	1	59						
2	2	48						
3	3	72						
4	4	24						
5	5	50						
6	6	23						
7	7	36						
8	8	66	2	26.2	114	255	185	56
9	9	60	2	32.1	83	179	119.4	42
10	10	29	1	30	85	180	93.4	43
11	11	22	1	18.6	97	114	57.6	46
12	12	56	2	28	85	184	144.8	32
13	13	53	1	23.7	92	186	109.2	62
14	14	50	2	26.2	97	186	105.4	49

Selected: 0 Rows: 50/442 Page size: 50 Page: 1 /9

2.17.1 Export

2.17.1.1 Data Source

1. Select data rows and columns.
2. Data filter can be set.

MyBox Export - CSV - 942 - d:\tmp\mybox-data-6.5.9\data\DiabetesPrediction_en.csv

Data Row filter Target formats Target files Logs

Rows Selected Current page All pages

Columns   

<input type="checkbox"/> Table r...	<input type="checkbox"/> Data...	<input checked="" type="checkbox"/> age	<input checked="" type="checkbox"/> sex	<input checked="" type="checkbox"/> BMI(b...	<input checked="" type="checkbox"/> BP(ave...	<input type="checkbox"/> S1(blo...	<input type="checkbox"/> S2(blo...	<input type="checkbox"/> S3 +
<input type="checkbox"/> 1	1	59	2	32.1	101	157	93.2	38
<input type="checkbox"/> 2	2	48	1	21.6	87	183	103.2	70
<input checked="" type="checkbox"/> 3	3	72	2	30.5	93	156	93.6	41
<input checked="" type="checkbox"/> 4	4	24	1	25.3	84	198	131.4	40
<input checked="" type="checkbox"/> 5	5	50	1	23	101	192	125.4	52
<input checked="" type="checkbox"/> 6	6	23	1	22.6	89	139	64.8	61
<input type="checkbox"/> 7	7	36	2	22	90	160	99.6	50
<input type="checkbox"/> 8	8	66	2	26.2	114	255	185	56
<input type="checkbox"/> 9	9	60	2	32.1	83	179	119.4	42
<input checked="" type="checkbox"/> 10	10	29	1	30	85	180	93.4	43
<input checked="" type="checkbox"/> 11	11	22	1	18.6	97	114	57.6	46
<input type="checkbox"/> 12	12	56	2	28	85	184	144.8	32
<input checked="" type="checkbox"/> 13	13	53	1	23.7	92	186	109.2	62
<input type="checkbox"/> 14	14	50	2	26.2	97	186	105.4	49
<input type="checkbox"/> 15	15	61	1	24	91	202	115.4	72

Selected: 7 Rows: [1-50]50/442 Page size 50 Page 1 /9     

2.17.1.2 Formats of Targets

1. Select targets and their formats: csv, texts, excel, xml, json, html, pdf, MyBox Clipboard.
2. Set options of target files.
3. Split files in maximum lines.
4. Select whether include row numbers.
5. Select whether save date/time/era and numbers as columns' formats.

My Box Export - CSV - 942 - d:\tmp\mybox-data-6.5.9\data\DiabetesPrediction_en.csv

Data Row filter Target formats Target files Logs

Base CSV Excel Texts html json PDF

Formats  

CSV XLS TXT JSON XML HTML PDF 

Split files in maximum lines Not split

Row number

Save date/time/era and numbers as columns' formats

2.17.1.3 Target Files

1. Set target path.
2. Select whether open files after export.

My Box Export - CSV - 942 - d:\tmp\mybox-data-6.5.9\data\DiabetesPrediction_en.csv

Data Row filter Target formats Target files Logs

Target path d:\tmp\mybox-data-6.5.9\generated 

Append timestamp in file name

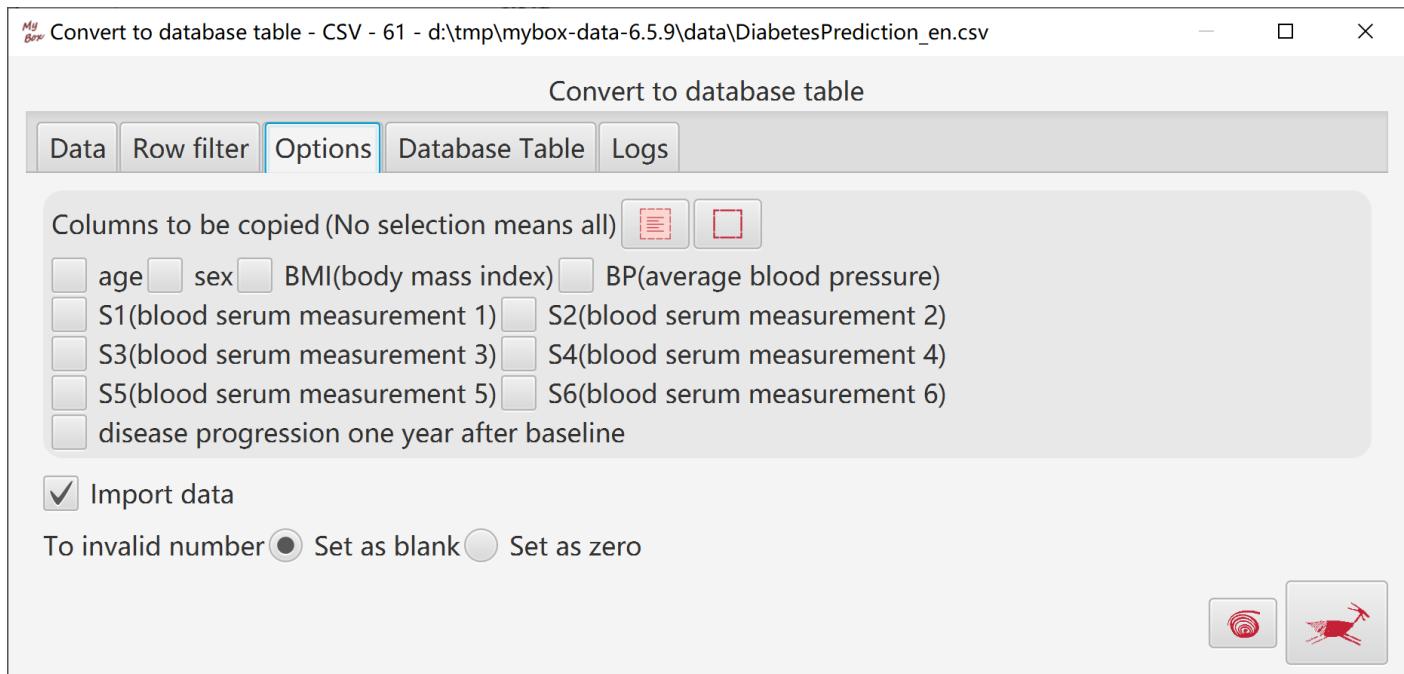
If files exist Replace Skip Target name append _m

Open files after export

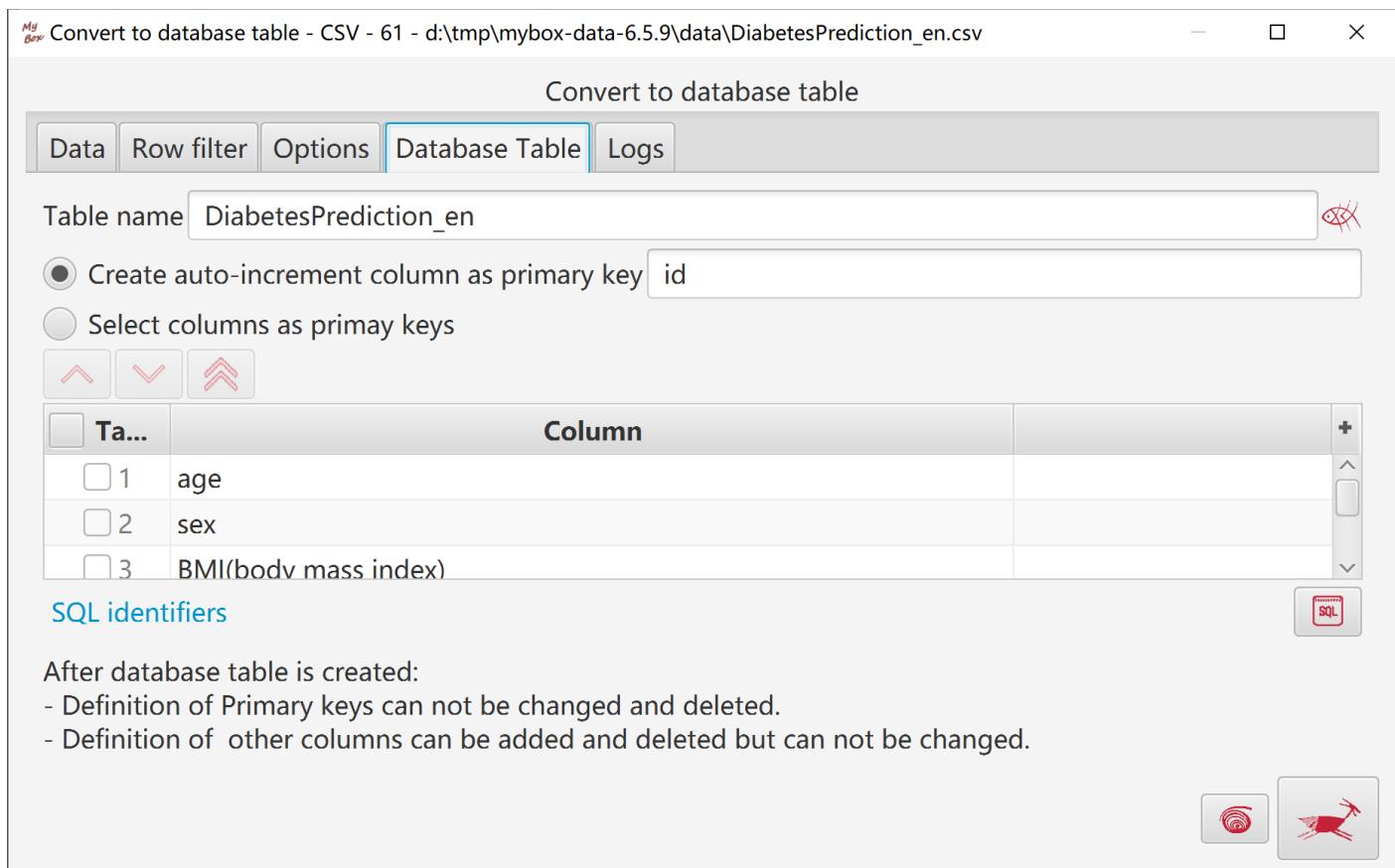
 

2.17.2 Convert to Database Table

1. Select data rows. Data filter can be set.
2. Select columns to be copied.
3. Select whether import data.



4. Create auto-increment column, or select some columns as primary key.



2.18 Manage Data

MyBox records definition of data objects in its internal table:

	csv/excel/text files	MyBox Clipboard	Matrice	Database tables
Time to create data definition	When data file is opened for the first time	When data are copied in MyBox Clipboard	When new matrix is saved in Matrices Manager	When new table is saved in Database Table Manager
Storage location of data	Data file	Files under MyBox internal path.	Database table of MyBox	Database tables of MyBox
When delete data definition	Data file is not affected	Internal file is deleted	Data of matrix are cleared	Database table is dropped

The screenshot shows the 'Manage Data' window for a database table named 'CHINAPOPULATION_EN'. The window has a toolbar at the top with various icons for file operations. Below the toolbar is a menu bar with options like Window, Document, Image, Data, File, Media, Network, Settings, Recent Accessed, Development, and Help. A sub-menu for 'File' is open, showing options like 'Open', 'Save', 'Import', etc. The main area contains a list of data objects on the left and a detailed table view on the right.

List of Data Objects:

- CSV Open
- Excel
- Texts
- Matrix
- Database Table
- MyBox Clipboard

Detailed View (Database Table CHINAPOPULATION_EN):

	Row...	Colu...	+ ↗	Data row	id	year_	populatio...	m + ↗
at...	73	6	↑	1	1	1949	54,167	28
at...	6	7	↑	2	2	1950	55,196	28
at...	6	7	↑	3	3	1951	56,300	29
at...	6	7	↑	4	4	1952	57,482	29
at...	1	7	↑	5	5	1953	58,796	30
at...	6	7	↑	6	6	1954	60,266	31

Query Options:

- Type: CSV, XLS, TXT, Matrix, Database Table, MyBox Clipboard, All
- Order by: Descending (radio button), Modify time (radio button selected), Name, ID, Rows number, Columns number, File

Pagination and Selection:

- Page: 1 / 10
- Page size: 50
- Selected: 0 Rows: 50/73 Page size: 50
- Page: 1 / 2

2.19 Splice Data

1. Select or open two data.
 2. Select rows and columns from the two data:
 - Rows can be: current page, selected rows, or all pages.
 - Select columns. If no column is selected, then all columns are taken.
 - Set row filter.
 3. Options:
 - Direction: vertical, horizontal
 - Rows/Columns number by: Data A, Data B, longer, shorter.
 4. Target can be following: new csv/excel/text file, matrix, system clipboard, MyBox clipboard, database table.

My Box Splice Data : Database Table - 1081 - CHINAPOPULATION_EN

Window Document Image Data File Media Network Settings Recent Accessed Development Help

Set as data B

Tab...	ID	Type		Colu...	
<input type="checkbox"/> 1	1105	Texts	ChinaPo...	73	6
<input type="checkbox"/> 2	1061	Texts	ChinaPo...	73	6
<input type="checkbox"/> 3	1089	MyBox Clipbo...	ChinaPo...	6	7
<input type="checkbox"/> 4	1086	MyBox Clipbo...	ChinaPo...	6	7
<input type="checkbox"/> 5	1085	MyBox Clipbo...	ChinaPo...	6	7
<input type="checkbox"/> 6	1083	MyBox Clipbo...	ChinaPo...	1	7
<input type="checkbox"/> 7	1082	MyBox Clipbo...	ChinaPo...	6	7
<input type="checkbox"/> 8	1063	MyBox Clipbo...	a	6	3
<input type="checkbox"/> 9	1065	Matrix	a	3	3
<input type="checkbox"/> 10	1064	Matrix	b	6	4
<input type="checkbox"/> 11	1062	Excel	ChinaPo...	73	6
<input checked="" type="checkbox"/> 12	1081	Database Table	ChinaPo...	73	7
<input type="checkbox"/> 13	1069	Database Table	ed	3	3
<input type="checkbox"/> 14	1068	Database Table	ChinaPo...	73	7
<input type="checkbox"/> 15	1103	CSV	ChinaPo...	73	6

Page 1 /10 Page size

Rows: 50/457 Selected: 1

Data A Data B Splice Data

Direction: Horizontal

Rows number: By Data A

Target: CSV

Name:

2.20 Data File

2.20.1 CSV File

1. In the file:
 - In general, the first line(header) defines column names, and each of followed lines defines a row of data.
 - Values are separated by "delimiter" which can be string.
 - If a value contains delimiter or line break, it should be surrounded by quotes.
 - If delimiter is “#”, lines started with “#” are skipped and treated as comments.
2. When the file is opened for the first time, the tool guesses its delimiter and charset.
3. When file is loaded abnormally, change options and click Refresh button.
4. Data are paginated. When pages number is larger than 1, changes should be saved before run some functions.
Options: charset, whether has first line as field names, and delimiter of data.
5. To string values, multiple lines can be edited and saved.
6. Data can be saved as different charsets and delimiters.

MyBox Edit CSV File : CSV - 61 - d:\tmp\mybox-data-6.5.9\data\DiabetesPrediction_en.csv

Window Document Image Data File Media Network Settings Recent Accessed Development Help

Information

Format

Charset

- Determine automatically
- Known

First line defines columns names

Text delimiter

- String
- Tab
- Blank
- Blank characters
- 4 blanks
- 8 blanks
- ,
- |
- #
-
- +
- :
- ;
- @
- &
- %
- !
- "
- '
- ?
- .
- *
- \
- /
- _
- =
- <
- >

If file is loaded incorrectly, change options and click button "Refresh"

Backup

Save as

Charset

First line defines columns names

Text delimiter

- String

Ta...	Data ...	age	sex
1	1	59	2
2	2	48	1
3	3	72	2
4	4	24	1
5	5	50	1
6	6	23	1
7	7	36	2
8	8	66	2
9	9	60	2
10	10	29	1
11	11	22	1
12	12	56	2
13	13	53	1
14	14	50	2
15	61	1	

Selected: 0 Rows: 50/442 Page size 50

Page 1 /9

2.20.2 Text File

1. In the file:
 - In general, the first line(header) defines column names, and each of followed lines defines a row of data.
 - Values are separated by "delimiter" which can be string.
Regular expression is supported when parse the file.
 - Values should not contain delimiter.
 - If a line starts with "#", it will be skipped.
2. When the file is opened for the first time, the tool guesses its delimiter and charset.
3. When file is loaded abnormally, change options and click Refresh button.
4. Data are paginated. When pages number is larger than 1, changes should be saved before run some functions.
Options: charset, whether has first line as field names, and delimiter of data.
5. Not support multiple lines in values.
6. Data can be saved as different charsets and delimiters.

MyBox Edit Text Data File : Texts - 69 - d:\tmp\mybox-data-6.5.9\data\ChinaCPI_en.csv

Window Document Image Data File Media Network Settings Recent Accessed Development Help

Information Format

Charset

Determine automatically

Known **UTF-8**

First line defines columns names

Text delimiter

String

Tab Blank Blank characters 4 blanks

8 blanks

, | # - + : ; @ & % !

" ' ? . * \ / _ = < >

If file is loaded incorrectly, change options and click button "Refresh"

Backup Save as

Charset **UTF-8**

First line defines columns names

Text delimiter

String

Ta...	Data ...	year_	consumer ...	urba ...
1	1	1951	112.5	
2	2	1952	102.7	
3	3	1953	105.1	
4	4	1954	101.4	
5	5	1955	100.3	
6	6	1956	99.9	
7	7	1957	102.6	
8	8	1958	98.9	
9	9	1959	100.3	
10	10	1960	102.5	
11	11	1961	116.1	
12	12	1962	103.8	
13	13	1963	94.1	
14	14	1964	96.3	
15	15	1965	98.8	

Selected: 0 Rows: 50/71 Page size 50 Page

1 /2

2.20.3 Excel File

1. Choose one worksheet to handle.
2. Worksheets can be added/renamed/deleted.
3. In general, the first line(header) defines column names, and each of followed lines defines a row of data.
4. If file is read abnormally, change options and click button "Refresh".
5. Data are paginated. When pages number is larger than 1, changes should be saved before run some functions.
6. To string values, multiple lines can be edited and saved.
7. Data can be saved with current sheet only or all worksheets.

Notice: Tool can only handle base data in Excel file.

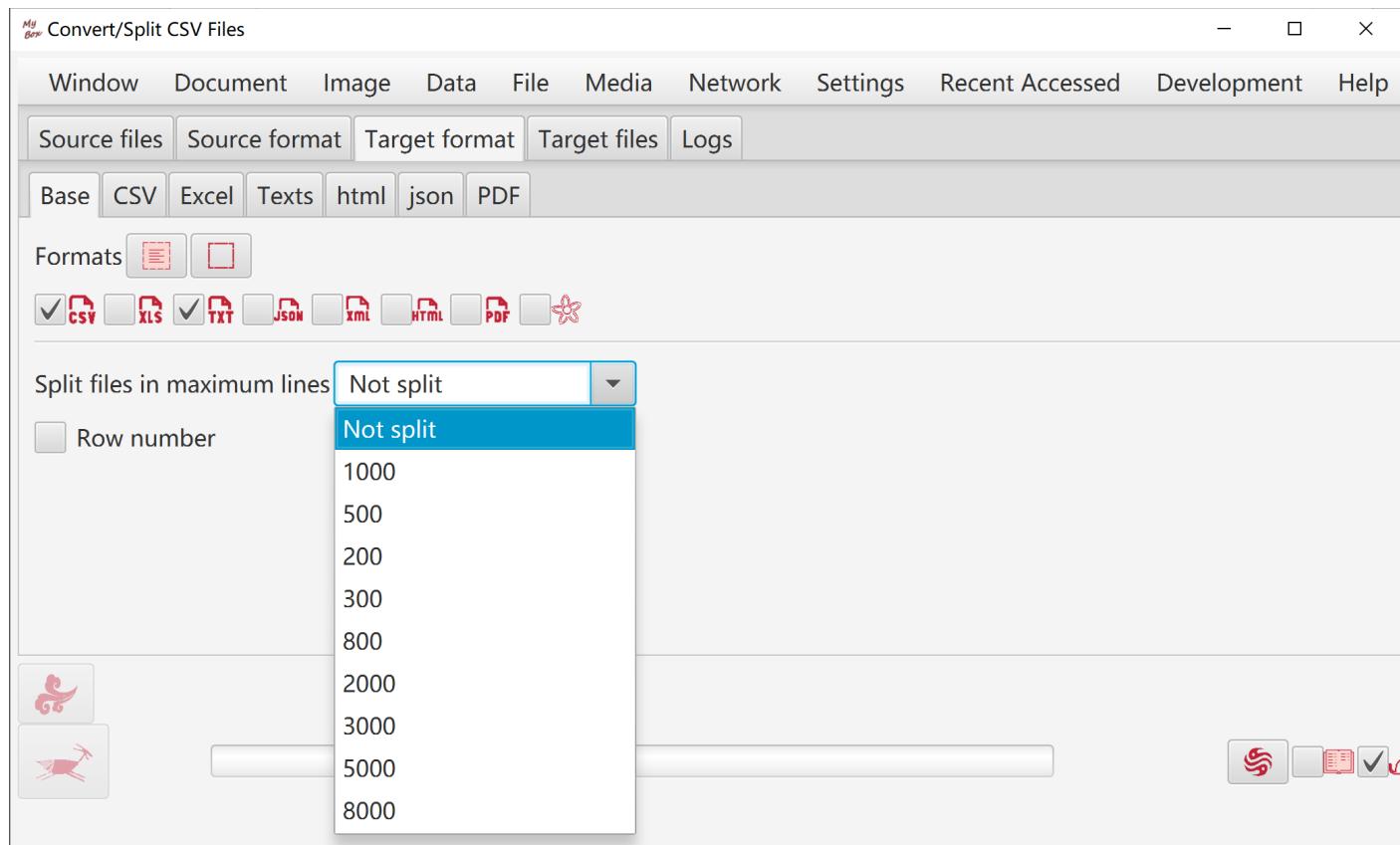
If file includes format, style, formula, or chart, suggest to save changes as new file to avoid data loss.

Tab...	Data row	year_	college gr...	middle sc...	high : .
1	1	1949	2.1	28	6.1
2	2	1950	1.8	29.6	6.2
3	3	1951	1.9	28.4	5.9
4	4	1952	3.2	22.1	3.6
5	5	1953	4.8	45.4	5.6
6	6	1954	4.7	64.4	6.8
7	7	1955	5.5	96.9	9.9
8	8	1956	6.3	93.9	15.4
9	9	1957	5.6	129.9	18.7
10	10	1958	7.2	131.3	19.7
11	11	1959	7	179	29.9
12	12	1960	13.6	171	28.8
13	13	1961	15.1	227.1	37.9
14	14	1962	17.7	202.5	44.1
15	15	1963	19.9	195.6	43.3
16	16	1964	20.4	175.3	36.7

Selected: 0 Rows: 20/73 Page size: 20 Page: 1 /4

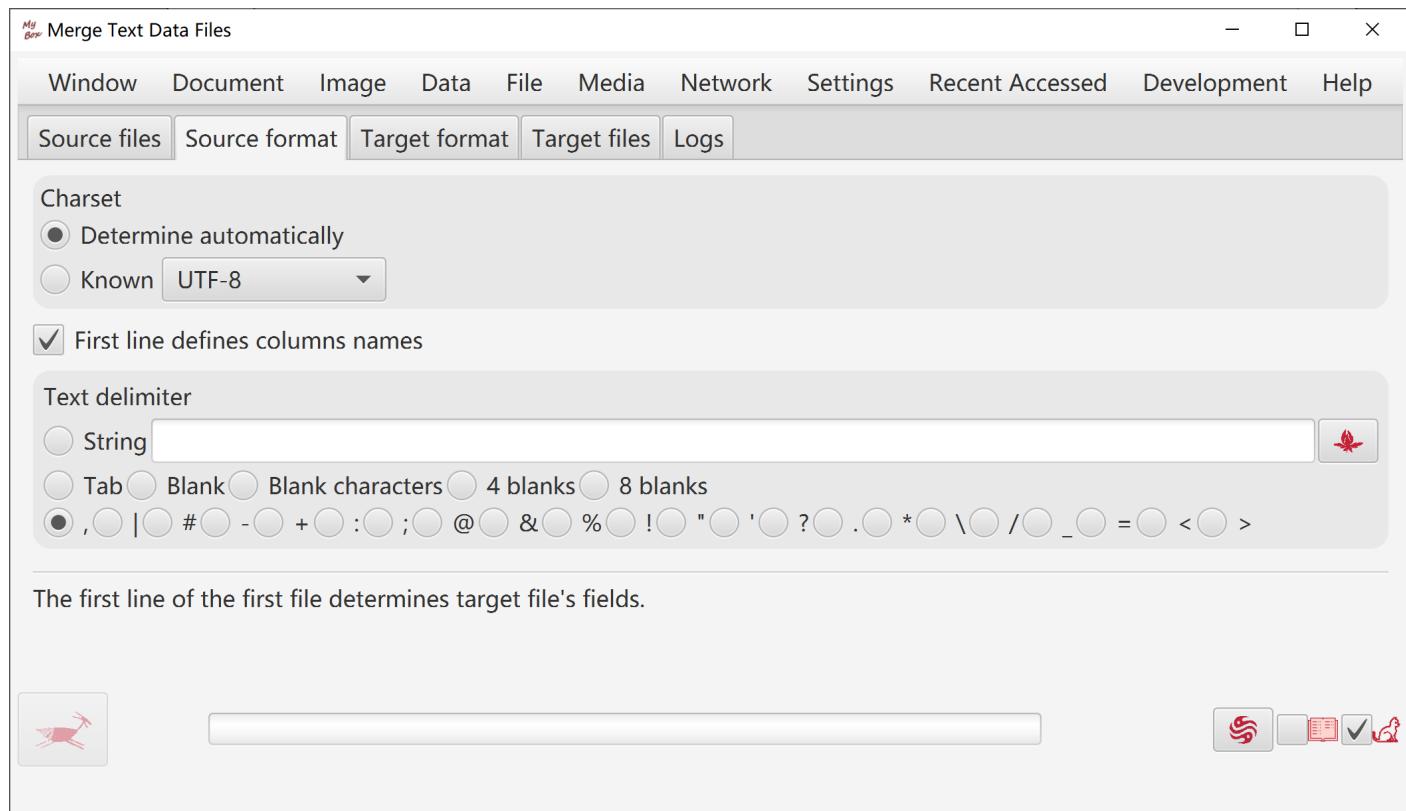
2.20.4 Convert/Split Data Files in Batch

1. Source files' formats can be csv, excel, and text. Options of source files can be set.
2. Target files' formats include csv, text, excel, xml, json, html, pdf. Options of target files can be set.
3. Split files as maximum lines.



2.20.5 Merge Data Files

1. Set source format.
2. Set target format.



2.21 Data in System Clipboard

1. Read contents in System Clipboard.
2. Guess delimiter and parse data.
3. User can select a delimiter to parse data.
4. Select how to parse:
 - CSV parser: Values can contain delimiter and line break. But not support parsing “Blanks” and “Regular Expression”.
 - Text parser: Values can not contain delimiter and line break. Support parsing “Blanks” and “Regular Expression”.
5. Select whether first line defines column names.

Data in System Clipboard

Window Document Image Data File Media Network Settings Recent Accessed Development Help

CSV Texts

First line defines the columns' names

Delimiter: ,

year_,population at year-end(ten thousand),male(ten thousand),female(ten thousand),urban(%)
 1949,54167,28145,26022,5765,48402
 1950,55196,28669,26527,6169,49027
 1951,56300,29231,27069,6632,49668
 1952,57482,29833,27649,7163,50319
 1953,58796,30468,28328,7826,50970
 1954,60266,31242,29024,8249,52017
 1955,61465,31809,29656,8285,53180
 1956,62828,32536,30292,9185,53643
 1957,64653,33469,31184,9949,54704
 1958,65994,34195,31799,10721,55273
 1959,67207,34890,32317,12371,54836
 1960,66207,34283,31924,13073,53134
 1961,65859,33880,31979,12707,53152
 1962,67296,34517,32778,11659,55636
 1963,69172,35533,33639,11646,57526
 1964,70499,36142,34357,12950,57549
 1965,72538,37128,35410,13045,59493
 1966,74542,38189,36353,13313,61229
 1967,76368,39115,37253,13548,62820
 1968,78534,40226,38308,13838,64696
 1969,80671,41289,39382,14117,66554
 1970,82992,42686,40306,14424,68568

If a value contains delimiter or line break, it should be surrounded by quotes.

Data...	year_	populatio...	male(ten t...)	female(te...)	urban(%)
1	1949	54167	28145	26022	5765
2	1950	55196	28669	26527	6169
3	1951	56300	29231	27069	6632
4	1952	57482	29833	27649	7163
5	1953	58796	30468	28328	7826
6	1954	60266	31242	29024	8249
7	1955	61465	31809	29656	8285
8	1956	62828	32536	30292	9185
9	1957	64653	33469	31184	9949
10	1958	65994	34195	31799	10721
11	1959	67207	34890	32317	12371
12	1960	66207	34283	31924	13073
13	1961	65859	33880	31979	12707
14	1962	67296	34517	32778	11659
15	1963	69172	35533	33639	11646
16	1964	70499	36142	34357	12950

Selected: 0 Rows: [1-20]20/50 Page size 20 Page 1 /3

2.22 Data in MyBox Clipboard

The screenshot shows the MyBox Data Tools application window titled "Data in MyBox Clipboard : MyBox Clipboard - 6 - b". The window has a menu bar with "Window", "Document", "Image", "Data", "File", "Media", "Network", "Settings", "Recent Accessed", "Development", and "Help". Below the menu is a toolbar with various icons.

The main area contains two tables. On the left, a table lists 238 rows of data with columns: Ta..., ID, Type, Name, Row..., Colu..., and a file path column. Row 7 is selected, showing "b" in the Name column. A modal dialog box is open over this table, prompting for a new name, with "bm" entered in the "New name" field. Buttons for "确定" (Confirm) and "取消" (Cancel) are at the bottom of the dialog.

On the right, there is another table view with columns: Tab..., Data row, 收入 (Income), 快乐 (Joy), and several other columns with numerical values. This table also has a "Selected: 0" status indicator at the bottom.

At the bottom of the interface, there are page navigation controls (Page 1 / 5, Previous, Next, Last), a "Page size" dropdown set to 50, and a "Rows: 7/7 Selected: 1" status bar.

2.23 Matrix

2.23.1 Edit and Manage Matrices

1. Edit matrix.
2. Matrix can be saved and reused.

The screenshot shows the 'Manage Matrices' window in MyBox Data Tools. The left panel lists matrices with columns for ID, Type, Name, Row..., and Colu... (with a plus sign). Matrix 1065 is selected, showing details for matrix 'a' (3 rows, 3 columns). The right panel shows the data for matrix 'a' in a table format with columns for Data row, 列1, 列2, and 列3. The data is as follows:

Data row	列1	列2	列3
1	353.63	469.35	59.21
2	687.2	344.24	308.21
3	359.8	359.52	925.57

Page navigation controls at the bottom indicate page 1 of 1, with 50 rows selected.

2.23.2 Unary Matrix Calculation

Transpose, Row Echelon Form, Reduced Row Echelon Form, Determinant By Elimination, Determinant By Complement Minor, Inverse Matrix By Elimination, Inverse Matrix By Adjoint, Matrix Rank, Adjoint Matrix, Complement Minor, Normalize, Multiply Number, Divide By Number, Power.

The screenshot shows the 'Matrix Unary Calculation' window. On the left, there is a table listing two matrices:

Ta...	ID	Type	Name	Row...	Colu...
<input checked="" type="checkbox"/> 1	1065	Matrix	a	3	3
<input type="checkbox"/> 2	1064	Matrix	b	6	4

On the right, there are several tabs and configuration options:

- Data** tab is selected.
- Calculation** tab is also visible.
- Result** tab is visible.
- Operations:** Transpose, Row echelon form, Reduced row echelon form, Complement minor, Normalize (selected), Multiply number, Divide number, Determinant by elimination, Determinant by complement minor, Inverse matrix by elimination, Inverse matrix by adjoint, Rank of matrix, Adjoint matrix, Power.
- According to:** Columns (selected), Rows, All.
- Algorithm:** MinMax (selected), L1(Sum), L2(Z-Score).
- Range:** -1.0, 1.0, (0,1), (-1,1), (-1.0,1.0).
- Calculate ENTER** button is highlighted with a black box.

At the bottom left, there are page navigation controls: Page, Page size, and a status bar showing 'Page 1 /1' and 'Rows: 2/2 Selected: 1'.

2.23.3 Binary Matrices Calculation

Plus, Minus, Hadamard Product, Kronecker Product, Horizontally Merge, Vertically Merge.

Matrices Binary Calculation : Matrix - New data

Window Document Image Data File Media Network Settings Recent Accessed Development Help

Ta...	ID	Name	Row
<input type="checkbox"/> 1	1065	a	3
<input type="checkbox"/> 2	1064	b	6

Matrix A Matrix B Calculation Result

Vertical merge Horizontal merge Plus Minus Multiply
 Hadamard product Kronecker product

Example:
 $A = \begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{21} \end{matrix}$

$B = \begin{matrix} b_{11} & b_{12} \\ b_{21} & b_{21} \end{matrix}$

Kronecker Product =
 $\begin{matrix} a_{11}*b_{11} & a_{11}*b_{12} & a_{12}*b_{11} & a_{12}*b_{12} \\ a_{11}*b_{21} & a_{11}*b_{21} & a_{12}*b_{21} & a_{12}*b_{21} \\ a_{21}*b_{11} & a_{21}*b_{12} & a_{22}*b_{11} & a_{22}*b_{12} \\ a_{21}*b_{21} & a_{21}*b_{21} & a_{22}*b_{21} & a_{22}*b_{21} \end{matrix}$

Page 1 /1
Page size 50 Rows: 2/2
Selected: 0

2.24 Database Tables

2.24.1 Manage Database Tables

1. View table definition.
2. Execute SQL.

Database Table : Database Table - 1113 - INCOMEHAPPINESS_EN

Window Document Image Data File Media Network Settings Recent Accessed Development Help

Columns Attributes View Table definition

Ta...	ID	Type	Name	Row...	Colu...	
<input checked="" type="checkbox"/>	1	1113	Databas...	IncomeHappi...	477	3
<input type="checkbox"/>	2	1081	Databas...	ChinaPopulat...	73	7
<input type="checkbox"/>	3	1069	Databas...	ed	3	3
<input type="checkbox"/>	4	1068	Databas...	ChinaPopulat...	73	7
<input type="checkbox"/>	5					
<input type="checkbox"/>	6					
<input type="checkbox"/>	7					
<input type="checkbox"/>	8					

INCOMEHAPPINESS_EN

Column	Type	Length	Not null	Primary key	Auto generated	Refer to table	Refer to column
id	Long	19	Yes	Yes	Yes		
income	Double	52					
happiness	Double	52					

```
CREATE TABLE INCOMEHAPPINESS_EN (
    id BIGINT NOT NULL GENERATED BY DEFAULT AS IDENTITY (START WITH 1, INCREMENT BY 1),
    income DOUBLE,
    happiness DOUBLE,
    PRIMARY KEY ( id )
)
```

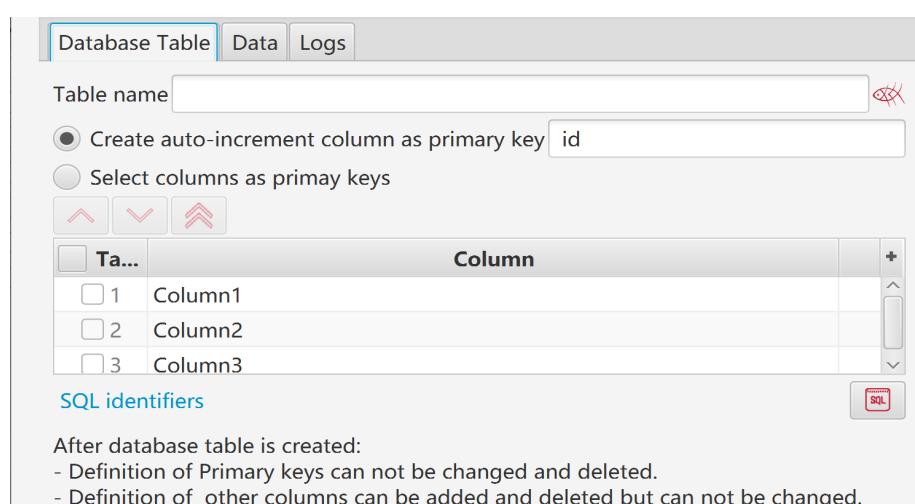
Page 1 /10

2.24.2 Limitations of SQL Identifier

1. Table name and column names should satisfy "Limitations of SQL identifier":
 - Maximum length is 128.
 - "Ordinary identifier":
 - Not surrounded by double quotation marks.
 - Must begin with a letter.
 - Contains only letters, underscore characters (_), and digits.
 - Permits Unicode letters and digits.
 - Can not be reserved words.
 - It is converted as uppercase when saved in database.
 - It is case-insensitive when referred in SQL statement.

Example, AbC is same as abc and aBC.
 - "Delimited identifier":
 - Surrounded by double quotation marks.
 - Can contain any characters.
 - It is saved as string inside the double quotations in database.
 - It should be surrounded by double quotations when referred in SQL statement, except for following: It only includes upper case letters and underscores.

Example, "AbC" is different from AbC or "ABC" while "ABC" is same as ABC and abc.
2. When MyBox create name of table/column: The name will be quoted if it includes invalid character.
3. After database table is created:
 - Definition of Primary keys can not be changed and deleted.
 - Definition of other columns can be added and deleted but can not be changed.



2.24.3 Database SQL

1. Provide examples of SQL statements.
2. List names of all user tables automatically.
3. View table definitions of all user tables.
4. Display outputs of execution and results of query.
5. SQL codes can be organized as information of tree.
6. Can load or save as external files.

The screenshot shows the MyBox Database SQL interface. The left sidebar contains a tree view under the 'SQL' category, with 'Maximum rows' selected. The main area has two tabs: 'Attributes' (selected) and 'SQL'. The SQL tab contains a text input field with the query: 'SELECT * FROM visit_history FETCH FIRST 300 ROWS ONLY'. Below the input field is a 'Results' tab. A data grid table is displayed with the following data:

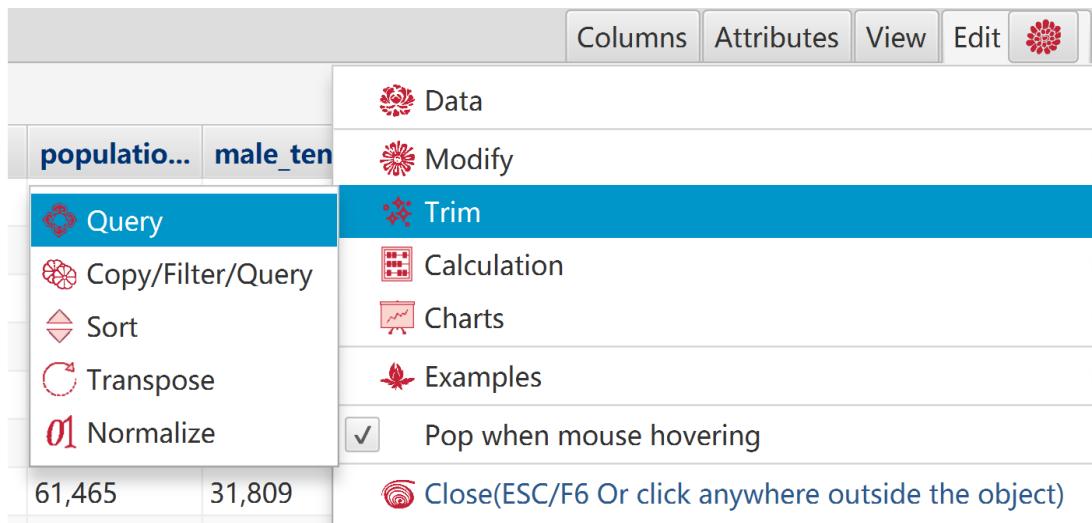
Data row	RESOURC...	FILE_TYPE	OPERATIO...	RESOURC...	DATA_MO...
1	2	21	3	d:\tmp\my...	20
2	1	21	3	d:\tmp\my...	20
3	2	21	2	d:\tmp\my...	20
4	1	21	2	d:\tmp\my...	20
5	2	21	3	d:\tmp\my...	20
6	1	21	3	d:\tmp\my...	20
7	2	21	2	d:\tmp\my...	20
8	1	21	2	d:\tmp\my...	20

At the bottom, there are navigation buttons for 'Selected: 0', 'Rows: 20/300', 'Page size: 20', 'Page: 1', '1/15', and arrows for navigating through the results.

2.24.4 SQL Query

Database table has a special function menu: “Functions” - “Trim” - “Query”, which can help to input and execute SQL query:

1. Names of table and columns are listed in left.
2. Provide examples and record histories.



Database Table : Database Table - 1115 - CHINAPOPULATION_EN *

Window Document Image Data File Media Network Settings Recent Accessed Development Help

Database Table - 1115 - CHINAPOPULATION_EN

Input a statement of SQL query which can be written in multiple lines

CHINAPOPULATION_EN	Attributes
id	Options
year_	SQL*
population_at_year_end_ten_thousand_	
male_ten_thousand_	
female_ten_thousand_	
urban_ten_thousand_	
rural_ten_thousand_	

SELECT year_, female_ten_thousand_, urban_ten_thousand_, rural_ten_thousand_ FROM CHINAPOPULATION_EN

3 Script and Expression

3.1 JShell(Java interactive coding tool)

3.1.1 About JShell

JShell is one of tools in JDK:

1. JShell provides capability to interactively evaluate "snippets", as Read-Eval-Print Loop (REPL).
2. "Snippet" is a single expression, statement, or declaration of Java programming language code:
 - Semicolons should be in the end of statement while expression need not it.
 - Variables and methods can be defined and called later.
3. External Java classes should be accessible:
 - JShell picks "CLASSPATH" of system environment.
 - Other jar files or paths can be appended to "CLASSPATH".
 - Except for base classes, most of Java classes should be imported before call them.
4. JShell can be used for scientific computation and Java codes debug.

3.1.2 Run Jshell in GUI

This tool helps to run JShell in GUI:

1. Input several snippets and click button "Start" to run them.
2. Snippets are evaluated one by one.
3. Results of snippets will affect later snippets, like "an execution environment".
4. Attributes of all evaluated snippets will be shown in a table.
5. Click button "Delete" or "Clear" to drop some or all snippets from current environment.
6. Click button "Reset" to empty JShell and environment becomes blank.
7. Press "CTRL+1" to pop list of code completion suggestions.
8. If added MyBox class paths, all methods of MyBox can be referred.
9. JShell codes are organized in tree. Examples are provided.

MyBox JShell(Java interactive coding tool): 911 - Format number

Window Document Image Data File Media Network Settings Recent Accessed Development Help

Nodes Tags Time Find Select Attributes Codes

```

import java.math.BigDecimal;
import java.math.RoundingMode;
double scale(double v, int scale) {
    BigDecimal b = new BigDecimal(v);
    return b.setScale(scale, RoundingMode.HALF_UP).d
}

import java.text.DecimalFormat;
String formatDouble(double data, int scale) {
    try {
        String format = "#,###";
        if (scale > 0) {
            format += "." + "#".repeat(scale);
        }
        import java.math.BigDecimal;
        import java.math.RoundingMode;
        double scale(double v, int scale) {
            BigDecimal b = new BigDecimal(v);
            return b.setScale(scale, RoundingMode.HALF_UP).doubleValue();
        }
    }
    import java.
    double circleAreaByRadius(double radius) {
        return Math.PI * radius * radius ;
    }

    formatDouble(circleAreaByRadius(273.4), 4)
}

```

Results Snippets Class paths

id: 18
Status: Valid
Type: METHOD
Name: formatDouble

2022-08-31 14:08:48
double circleAreaByRadius(double radius) {
 return Math.PI * radius * radius ;
}

id: 19
Status: Valid
Type: METHOD
Name: circleAreaByRadius

2022-08-31 14:08:48
double circleAreaByRadius(double radius) {
 return Math.PI * radius * radius ;
}

 id: 20
Status: Valid
Type: VAR
Name: \$9
Value: "234,826.3854"

Loaded

MyBox JShell(Java interactive coding tool): 911 - Format number

Window Document Image Data File Media Network Settings Recent Accessed Development Help

Nodes Tags Time Find Select Attributes Codes

```

import java.math.BigDecimal;
import java.math.RoundingMode;
double scale(double v, int scale) {
    BigDecimal b = new BigDecimal(v);
    return b.setScale(scale, RoundingMode.HALF_UP).d
}

import java.text.DecimalFormat;
String formatDouble(double data, int scale) {
    try {
        String format = "#,###";
        if (scale > 0) {
            format += "." + "#".repeat(scale);
        }
        import java.math.BigDecimal;
        import java.math.RoundingMode;
        double scale(double v, int scale) {
            BigDecimal b = new BigDecimal(v);
            return b.setScale(scale, RoundingMode.HALF_UP).doubleValue();
        }
    }
    import java.
    double circleAreaByRadius(double radius) {
        return Math.PI * radius * radius ;
    }

    formatDouble(circleAreaByRadius(273.4), 4)
}

```

Results Snippets Class paths

id: 18
Status: Valid
Type: METHOD
Name: formatDouble

2022-08-31 14:08:48
double circleAreaByRadius(double radius) {
 return Math.PI * radius * radius ;
}

id: 19
Status: Valid
Type: METHOD
Name: circleAreaByRadius

2022-08-31 14:08:48
double circleAreaByRadius(double radius) {
 return Math.PI * radius * radius ;
}

 id: 20
Status: Valid
Type: VAR
Name: \$9
Value: "234,826.3854"

Loaded

3.2 JEXL(Java Expression Language)

3.2.1 About JEXL

JEXL(Java Expression Language) is a library to generate values dynamically with variables and scripts.

1. JEXL has different syntax from Java. It is more like javascript.
2. Before run expression/script, all variables in it should have values held by JexlContext.
3. Refer to Java classes by creating their instances as local variables. Full package name is required.
4. JEXL can be used for scientific computation and data manufacture.

3.2.2 Run JEXL in GUI

This tool helps to run JEXL in GUI:

1. Input JEXL expression/script.

Notice: Use single quotes instead of double quotes to surround strings.

2. Input Java codes of setting JexlContext like following:

```
jexlContext.set("name", value);
```

Example, set following to use Math.PI in expression/script:

```
jexlContext.set("Math", Math.class);
```

3. Input parameters of JEXL script if any. Separate values by comma.

4. Click button "Start" to evaluate the expression/script.

5. MyBox does following in JShell environment automatically:

- Add MyBox library paths to CLASSPATH.
- Import necessary JEXL packages.
- Execute codes of JexlContext.
- Calculate expression/script with parameters(if any).

6. If all variables and parameters have valid values, result is shown in right pane.

7. JEXL codes are organized in tree. Examples are provided.

MyBox User Guide – Data Tools v6.7

The screenshot shows the MyBox Data Tools application window. The left sidebar contains a tree view of JEXL code categories: Examples, Object, Math, Boolean expression, Numeric operations, Date, and Statement. The 'percentage' node under 'Numeric operations' is selected and highlighted in blue. The main workspace is divided into several panels:

- JEXL script:** Displays the following code:

```
var number = 37;  
var total = 518;  
var scale = 2;  
DoubleTools.percentage(number, total, scale);
```
- JEXL Context:** Displays the following code:

```
jexlContext.set("DoubleTools", mara.mybox.tools.DoubleTools.class);
```
- Jexl script paramters:** An empty panel.
- Results:** A panel showing the execution history:
 - 2022-08-31 14:13:54: id: 29, Status: Valid, Type: STATEMENT. The code is the same as in the JEXL script panel.
 - 2022-08-31 14:13:54: id: 30, Status: Valid, Type: EXPRESSION, Name: jexlScript. The code is the same as in the JEXL Context panel.
 - 2022-08-31 14:13:54: id: 31, Status: Valid, Type: VAR, Name: \$7. The code is the same as in the JEXL Context panel.

3.3 Javascript

This tool helps to manage and run codes in JavaScript:

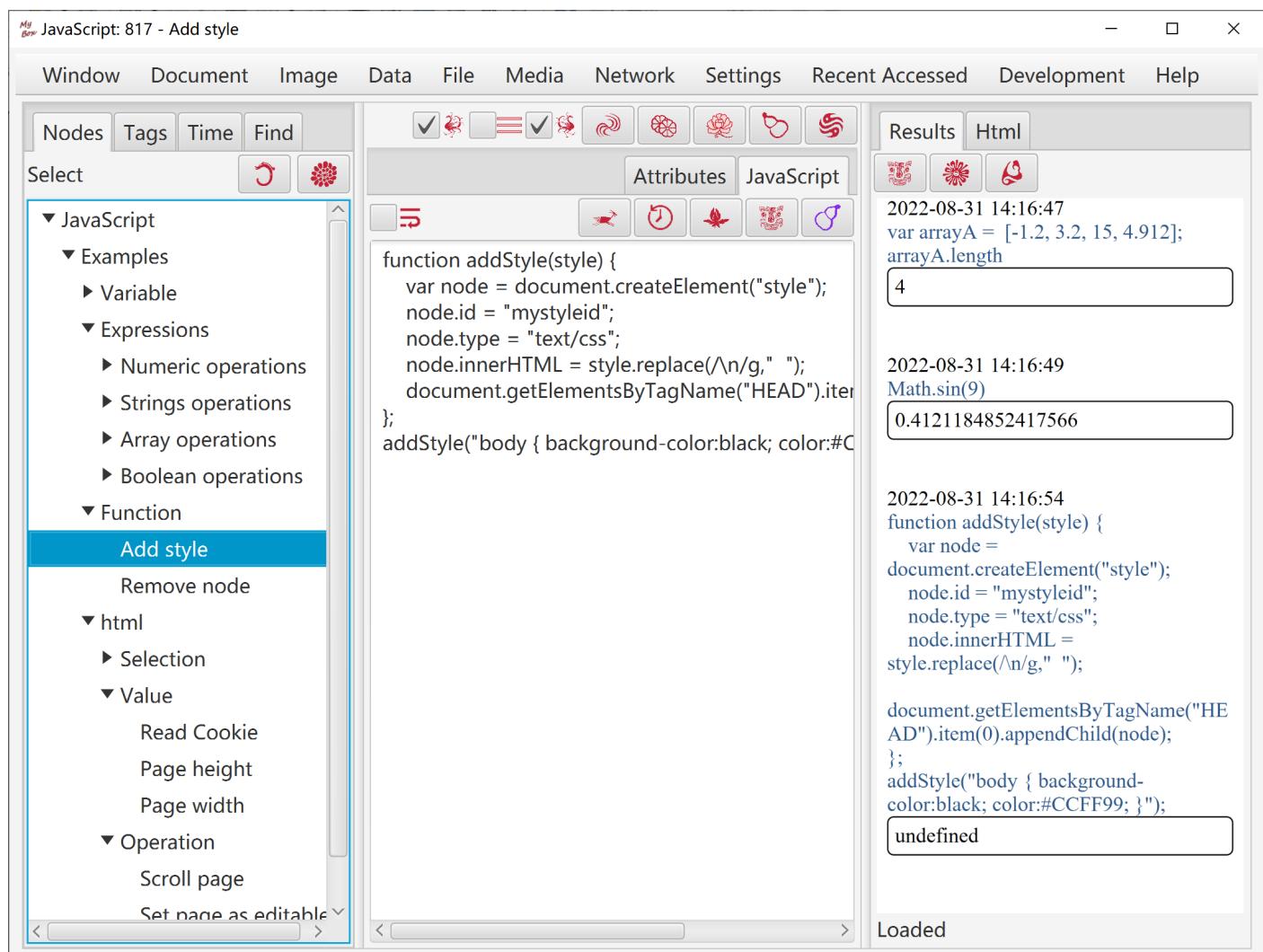
1. Edit codes in JavaScript.

The script can include any valid elements which WebEngine can parse(ECMAScript 6).

2. Run the script.

Its results are displayed in right pane. And it affects the web page in right pane too.

3. Javascript codes are organized in tree. Examples are provided.



4 Math Function

This tool helps to manage and calculate math functions:

4.1 Define Math Functions

1. List names of variables of the function, separated by English comma.

The names should satisfy rules of JavaScript:

- Can include letters, number, underline '_', and unicodes(like Chinese).
- Not start with number.

2. Define function expression as a piece of codes in JavaScript:

- The script can include any valid elements which Nashorn can parse(ECMAScript 5.1).
- It should be a number finally.
- It can refer to the given variables, but should not declare them.

3. Define domain of the function as a piece of codes in JavaScript:

- Blank means the domain is all of real numbers.
- The script can include any valid elements which Nashorn can parse(ECMAScript 5.1).
- It should be a boolean value(true or false) finally.
- It can refer to the given variables, but should not declare them.

4. Give the name of calculation result.

5. Saved in tree.

6. Examples are provided.

4.2 Calculate Math Function

1. When the tool calculates a script, 'var <variable_name>=<variable_value>;' are inserted in the head of it before the evaluation.
2. If script of domain is not blank, the given values of variables are inserted in it and check the result:
 - If true, continue to calculate the script of function expression.
 - Else this set of values is skipped.

Math Function: 217 - univariate logarithmic

Window Document Image Data File Media Network Settings Recent Accessed Development Help

Nodes Tags Time Find Select

Math Function

- Examples
- Constants
- Unary Function
- Binary Function**
 - Univariate Polynomial
 - Geometry
 - Quadric Surface
 - Trigonometry
 - Exponential Function
 - Logarithmic Function**
 - univariate logarithmic
- Numerical Analysis
- Univariate Normal probability
- Univariate Standard normal pr
- Ternary Function
- Trigonometry
- Geometry
- area of ellipsoid - half_axie
- area of ellipsoid - axie
- volumn of ellipsoid
- volumn of ellipsoid - axie

Attributes Math Function
Names of variables, separated by English comma
x, y

Expression
5 * Math.log(x * y) - 9

Function domain
x * y > 0

Result name univariate logarithmic

Calculate Data Set
x y
Decimal scale 8

```
var x=4.0;
var y=5.0;
var xMean = 1;
var xStd = 2;
var yMean = 2;
var yStd = 1;
var coefficient = 0.2;
var dx = x - xMean;
var dy = y - yMean;
var xyStd = xStd * yStd;
var dco = 1 - coefficient * coefficient;
var px = dx * dx / ( xStd * xStd );
var py = dy * dy / ( yStd * yStd );
var pxy = 2 * coefficient * dx * dy /
var expP = - (px + py - pxy) / ( 2 * coefficient );
var div = 2 * Math.PI * xyStd * Math.sqrt(1 - dco);
Math.exp(expP) / div
```

Univariate Normal probability density function

2022-09-28 14:13:12
var angle=45.0;
angle * Math.PI / 180
radian=0.78539816

4.3 Data Set

With defined range, interval, and decimal scale, data set of the function can be generated.

The screenshot shows the 'Math Function' configuration window in MyBox Data Tools. The title bar reads "My Math Function: 225 - law of sines - for edge(radian)". The menu bar includes Window, Document, Image, Data, File, Media, Network, Settings, Recent Accessed, Development, and Help. The left sidebar has tabs for Nodes, Tags, Time, and Find, with 'Select' currently selected. A tree view under 'Math Function' shows categories like Examples, Constants, Unary Function, Trigonometry, Geometry, Power Function, Exponential Function, Logarithmic Function, Intercept Function, Unitary Polynomial, piecewise values, Numerical Analysis, Binary Function, Ternary Function, and Trigonometry. Under Trigonometry, 'law of sines - for edge(radian)' is selected and highlighted in blue. The main workspace contains several toolbars with icons for selection, attributes, and other functions. A central panel shows the expression `edge_a * Math.sin(angleB_radian) / Math.sin(angleA_radian)`. To the right, there are tabs for Calculate and Data Set. The Data Set tab is active, showing settings for variables: `angleA_radian`, `edge_a`, and `angleB_radian`. It includes a 'Range' section with 'From' set to -10.0 and 'To' set to 10.0, and a 'Number of split' option set to 50. There is also a 'Data interval' option set to 0.1. Below these is a 'Decimal scale' dropdown set to 2. At the bottom, there is a 'Function domain' section with the condition `angleA_radian > 0 && angleB_radian > 0 && edge_a > 0` and a result name input field set to `edge_b`.

4.4 XY Chart of Unary Function

To unary function, XY charts, including scatter chart and line chart, can be displayed.

The screenshot shows the MyBox Data Tools interface with the following components:

- Top Bar:** Window, Document, Image, Data, File, Media, Network, Settings, Recent Accessed, Development, Help.
- Left Sidebar (Select):**
 - Math Function:**
 - Examples
 - Constants
 - Unary Function** (selected)
 - absolute
 - Trigonometry
 - Geometry
 - Power Function
 - Exponential Function
 - Logarithmic Function
 - Intercept Function
 - Unitary Polynomial
 - piecewise values
 - direct values
 - define functions
 - Numerical Analysis
 - Unary Normal probability de
 - Unary Standard normal prob
 - sigmoid** (selected)
 - derivative of sigmoid
 - Binary Function
 - Univariate Polynomial
 - Geometry
 - Quadric Surface
- Middle Panel (Math Function):**
 - Attributes** tab (selected)
 - Math Function** tab
 - Names of variables, separated by English comma: `x`
 - Expression: `1 / (1 + Math.exp(-x))`
- Right Panel (Calculate):**
 - Calculate button
 - Data Set button
 - x** input field
 - Range settings: From -20.0, To 20.0
 - Number of split: 100 (radio button selected)
 - Data interval: 0.1 (radio button)
 - Decimal scale: 2
- Bottom Panel (XY chart - sigmoid):**
 - Scatter chart (radio button selected)
 - Line chart (radio button)
 - Plot area showing the sigmoid function (blue line) and data points (blue circles).
 - Result name: sigmoid

4.5 XYZ Chart of Binary Function

To binary function, XYZ chart, including 3D scatter chart and surface chart, can be displayed.

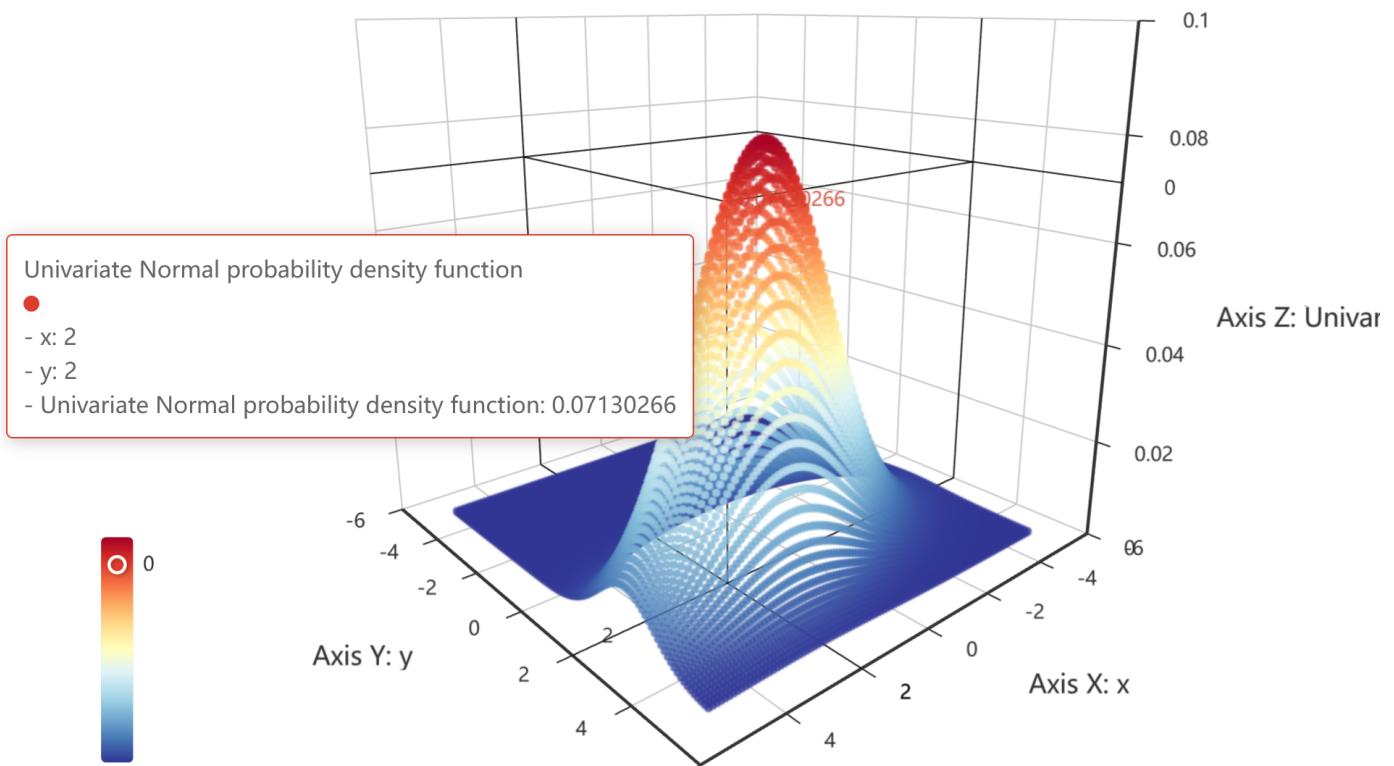
mybox-data-6.5.9/generated/Univariate_Normal_probability_density_function_Scatter_chart_2022-09-28_14-2 ☆



Univariate Normal probability density function - Scatter chart

Axis X: "x" Axis Y: "y"

Axis Z: ● Univariate Normal probability density function



5 Data of Location

5.1 Data Constraints

5.1.1 Invalid Value

1. Null value of integer/long/short is the minimum value(MIN_VALUE)
2. Null value of double is the maximum value(Double.MAX_VALUE)

5.1.2 Coordinate System:

1. CGCS2000(China Geodetic Coordinate System), real locations and approximate to WGS-84(GPS).
2. GCJ-02(China encrypted coordinate), encrypted data with offsets of real locations.
3. WGS-84(GPS), real locations.
4. BD-09(Baidu encrypted coordinate), based on GCJ-02.
5. Mapbar coordinate, based on GCJ-02.
6. When coordinate is unknown or invalid, the default value is CGCS2000.

5.1.3 Coordinate Values

1. Decimal values of longitude and latitude, instead of Degrees Minutes Seconds(DMS), are used when data handled.
2. MyBox provides "Location Tools" to convert coordinate values between decimal and DMS.
3. Valid range of longitude is '-180~180', and valid range of latitude is '-90~90'.

5.1.4 Time

5.1.4.1 Date Formats

- Date and Time, like: 2014-06-11 13:51:33
- Date, like: 2014-06-11
- Year, like: 2014
- Month, like: 2014-06
- Time, like: 13:51:33
- Time with Milliseconds, like: 13:51:33.261
- Date and Time with Milliseconds, like: 2014-06-11 13:51:33.261
- Date and Time with zone, like: 2020-09-27 12:29:29 +0800
- Date and Time with Milliseconds and zone, like: 2020-09-27 12:29:29.713 +0800
- "T" can be written or omitted between date and time. "2014-06-11T13:51:33" equals to "2014-06-11 13:51:33".

5.1.4.2 Era

"0 AD" = "1 BC" = "0" = "-0" = "0000" = "-0000" = "0001-01-01 00:00:00 BC" = "公元前 1" = "公元前

0001-01-01 00:00:00"

"1 AD" = "1" = "0001" = "0001-01-01 00:00:00" = "0001-01-01 00:00:00 AD" = "公元 1" = "公元 0001-01-01 00:00:00"

"202 BC" = "-203" = "-0203" = "-0203-01-01 00:00:00" = "0202-01-01 00:00:00 BC" = "公元前 202" = "公元前 0202-01-01 00:00:00"

"202 AD" = "202" = "0202" = "0202-01-01 00:00:00" = "0202-01-01 00:00:00 AD" = "公元 202" = "公元 0202-01-01 00:00:00"

5.1.4.3 Examples

2020-07-13 11:30:59

-2020-07-13 11:30:59

-581-01-23

960

公元 960

公元前 770-12-11

公元前 1046-03-10 10:10:10

202 BC

960-01-23 AD

1046-03-10 10:10:10 BC

5.2 Data Operations

1. Add/Delete/Edit/Copy/Clear/Refresh data.
2. Query data:
 - Define and manage query conditions.
 - Current query conditions is displayed on tab "information".
 - Data satisfying current query condition are displayed in tab "Data" in pages.
 - Data rows can be displayed in different colors as values of some column.
3. Import data in csv format:
 - File encoding is UTF-8 or ASCII.
 - The first line defines data headers which are delimited by English commas.
 - Followed each line defines one data row. Data fields are delimited by English commas.
 - The order of fields is not cared.
 - Necessary fields must occupy their locations, but need not have valid values(related to data).
 - Select whether replace existed data. Predefined data or example data always replace existed values.
4. Export data:
 - Define and manage export conditions.

- Export data fields can be selected.
 - Export file format can be selected: csv, xml, json, xlsx, html, pdf.
 - Select maximum lines to split files.
 - Can export current data page.
5. Delete/Clear data:
- Define and manage delete conditions.
 - Predefined data can not be deleted.
 - Referred data(like foreign keys) can not be deleted.
6. Define, manage, and use "Conditions":
- "Conditions" are used to execute querying, deleting, or exporting.
 - Set conditions in panes:
 - Data conditions are organized as trees. Multiple nodes can be selected.
 - Multiple data fields can be selected as sorting conditions, and their orders can be changed.
 - Edit condition: Title, where, order by, fetch. They will be merged as final conditon.
 - Manage conditons: Add/delete/edit/copy.
 - Conditions ever executed are saved automatically.
 - Recently visited conditions are listed in pop window of the buttons.

5.3 Map Data

1. Kinds of data can be presented in map, including Geography Codes, Location Data, and Coordinate Querying.
2. Data in map can be:
 - All data which satisfy current query condition. "Maximum number of data" can be set to avoid performance issues.
 - Data in current page.
3. TianDiTu:
 - Accepts coordinates of CGCS2000 and display them at correct locations without offsets.
 - When display other coordinates, MyBox converts them to CGCS2000 to show correct locations.
 - Projection can be selected: EPSG:900913/3857(Web Mercator) or EPSG:4326(Geodetic).
 - Controls can selected: Zoom, Scale, Map Type, Symbols.
 - Map Types: Standard, Satellite, Mixed Satellite, Terrain, Mixed Terrain.
 - Languages in different regions.
 - Range of map levels is 1-18.
4. GaoDe Map:
 - Accepts coordinates of GCJ-02 and display them at correct locations without offsets.

- When display other coordinates, MyBox converts them to GCJ-02 to show correct locations.
 - Projection is EPSG:900913/3857(Web Mercator).
 - Map layers:
 - Can select multiples: standard, satellite, roadnet, traffic.
 - Roadnet layer and traffic layer are only supported for China.
 - Satellite layer is supported for part of foreign addresses.
 - Opacity can be set for each map layer.
 - Map language: Chinese, English, Chinese and English.
 - Range of map levels is 3-18
 - Can selected "Fit View" to adjust map level and center as best automatically while display all data.
5. Adjust map level by:
- Scroll mouse wheel.
 - Click map controls.
 - Select "Map Size"
6. Marker image:
- Selections: point(bubble), circle, or any image.
 - For Location Data, more selections: Data Set Image, Data Image. Point will be used if no valid value.
 - Size can be set(Same size for width and height)
7. Marker text:
- Selections: Label, Coordinate, Address.
 - For Location Data, more selections: Start Time, End Time, Data Value, etc.
 - Multiples selections can be picked. Each selection will be showns in a line.
 - Size can be set.
 - Can select whether text is bold.
 - Color can be set. For Location Data, "Data Color" can be chosen.
8. Pop information:
- Detailed information can be popped when mouse is upon marker.
 - Can select whether pop information.
9. Snapshot:
- DPI can be set.
 - Current map and data in map can be saved and displayed in html.
10. Keys of map can be changed in "Settings". The default keys are free and shared by all MyBox users.

5.4 Geography Code

5.4.1 Data Definition

1. Basical attributes: id, level, longitude, latitude, chinese_name, english_name, 5 codes, 5 aliases,
2. Subordinate: owner, continent, country, province, city, county, town, village, building. ("Ancestors")
3. Auxiliary attributes: altitude, precision, coordinate system, area(square meters), population, comments, isPredefined.

5.4.2 Data Constraints

1. Not null values: id, level, chinese_name or english_name.
2. Values of "level": global(only "Earth"), continent, country, province(state), city, county(district), town, village(neighborhood), building, point of interest.
3. Data is unnecessary to be subordinated level by level. Cross-over can happen. Example, a village is subordinated to Antarctica, and a city belongs to a country without province level.
4. Match data:
 - One of following can determine an address:
 - Match "id"(assigned by MyBox automatically). This is accurate matching.
 - Match "level" + ancestors + "chinese_name"/"english_name"/any one "alias". This is accurate matching.
 - Match "level" + "chinese_name"/"english_name"/any one "alias". This is fuzzy matching. Duplicated names in same level can cause false matching.
 - Matching of name or alias is case-insensitive.
 - Sometimes 5 "code" are useful to match data.

5.4.3 Edit Data

1. "subordinate" of data is set by selecting node in locations tree.
2. "level" of data should be lower than its ancestors.
3. Data must have either chinese_name or english_name.
4. Select or display coordinate in map.
5. Set as "Predefined data" or "Inputted data" against selected rows.

5.4.4 Define Condition

All geogahy codes in MyBox are organized as a Locations Tree by their subordination relationship. Multiple nodes can be selected.

5.4.5 Import Data

5.4.5.1 Predefined Data

Include continents, countries, Chinese provinces /cities /counties.

Countries have values of "area" and "population".

5.4.5.2 CSV Format

- Download address:
https://github.com/Mararsh/MyBox_data/tree/master/md/GeographyCode/en
- Necessary fields:
Level,Longitude,Latitude
And "Chinese Name" or "English Name"
- Optional fields:
Altitude,Precision,Coordinate System,Square Kilometers,Population,
Code 1,Code 2,Code 3,Code 4,Code 5,Alias 1,Alias 2,Alias 3,Alias 4,Alias 5,
Continent,Country,Province,City,County,Town,Village,Building,Comments

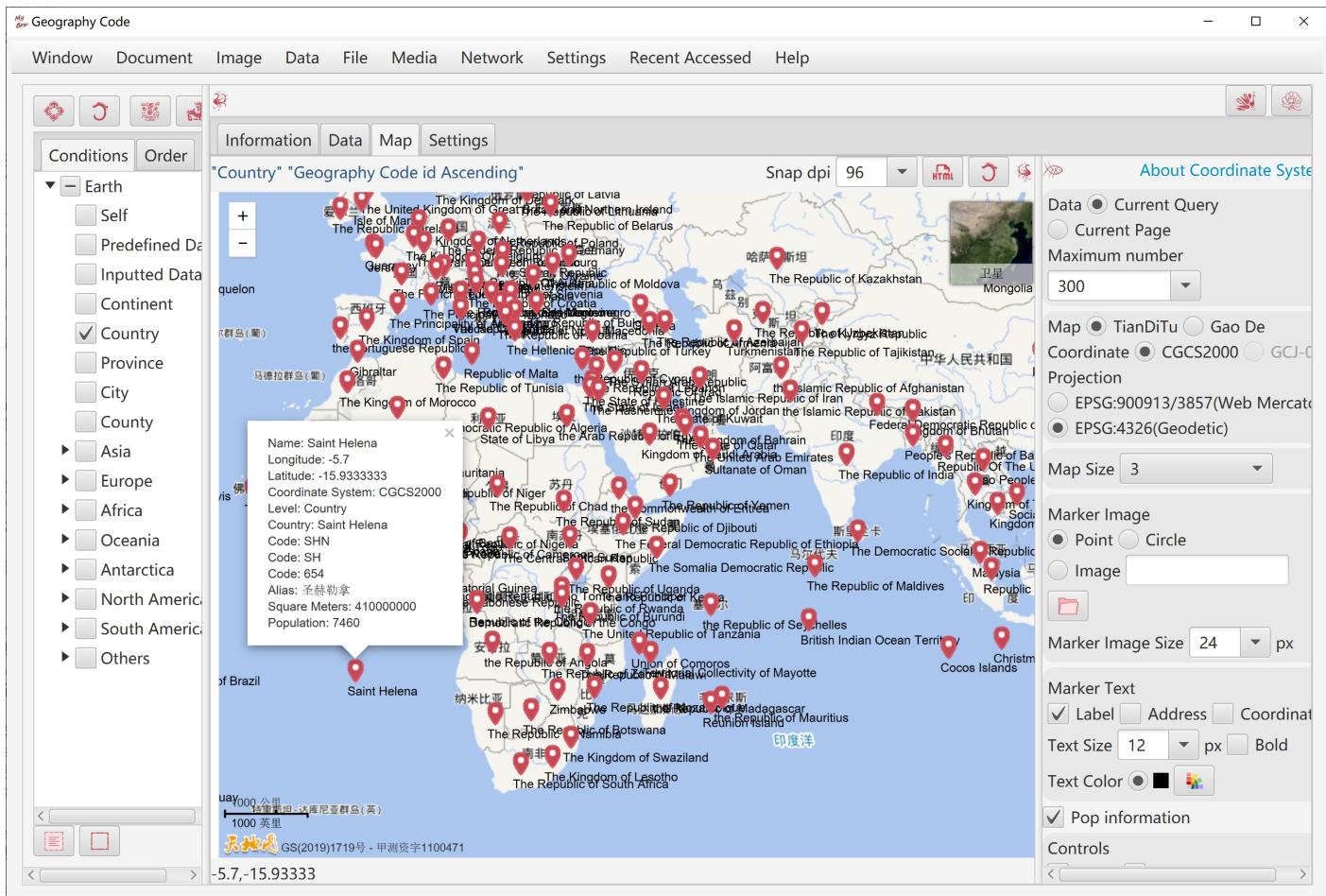
5.4.5.3 Data from geoname.org:

- Download address:
<http://download.geonames.org/export/zip/>
- Tab-delimited text in UTF8 encoding.
- Data fields:
countryCode postalCode placeName
adminName1 adminCode1 adminName2 adminCode2 adminName3 adminCode3
latitude longitude accuracy
- Coordinate system is WGS_84.
- Same address is written only once even when it has multiple "postal code" or coordinates.

5.4.6 Settings

1. Customize colors of data rows.
2. Provide "Default" and "Random" buttons.

MyBox User Guide – Data Tools v6.7



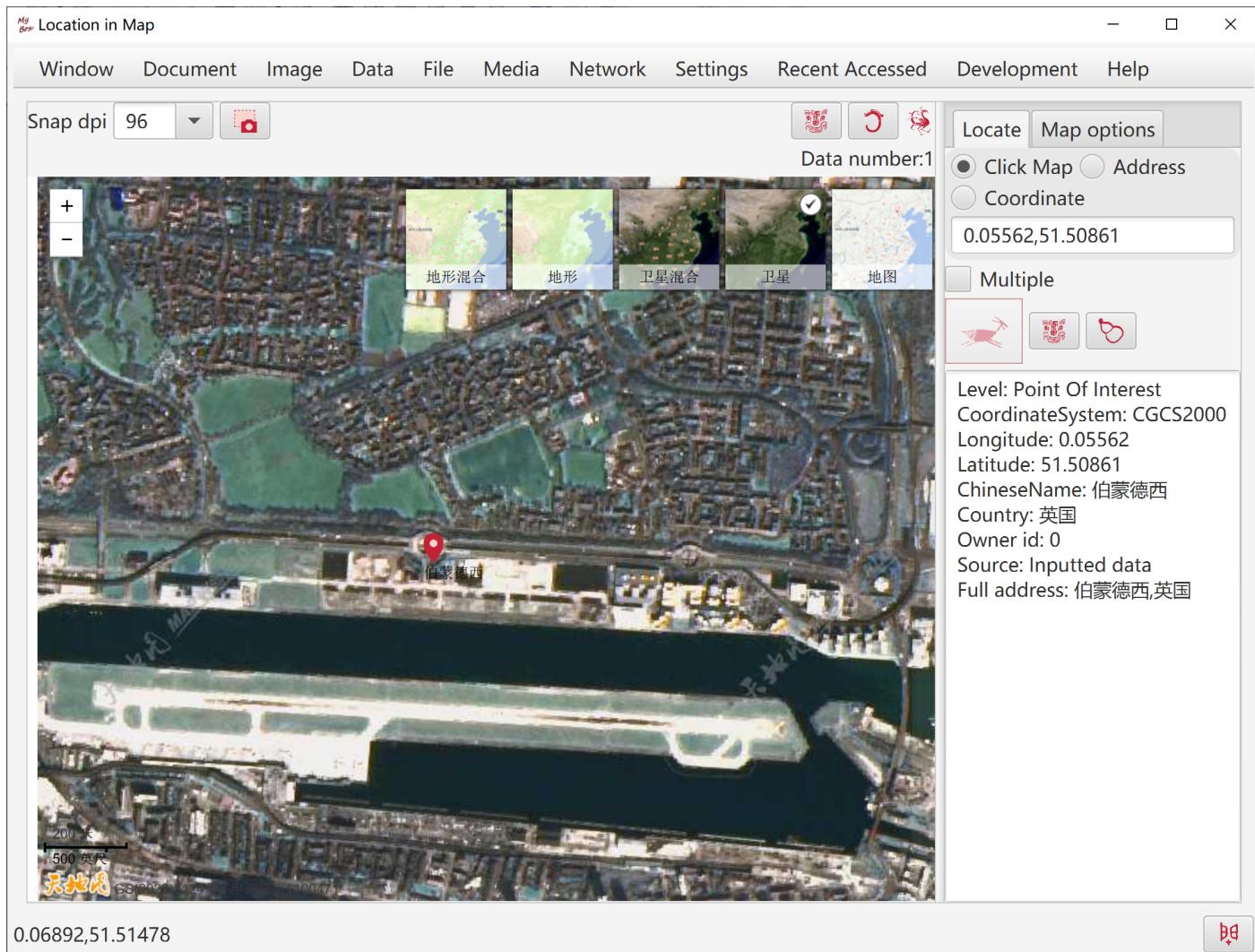
5.5 Location in Map

1. Query geography code by:

- Click map.
- Input address.
 - TianDiTu supports chinese and foreign addresses in Chinese(like "伦敦") or in English(like "Paris")
 - GaoDe map only supports addresses in China.

▪ Input longitude and latitude.

2. Query result can be saved in Geography Code table.

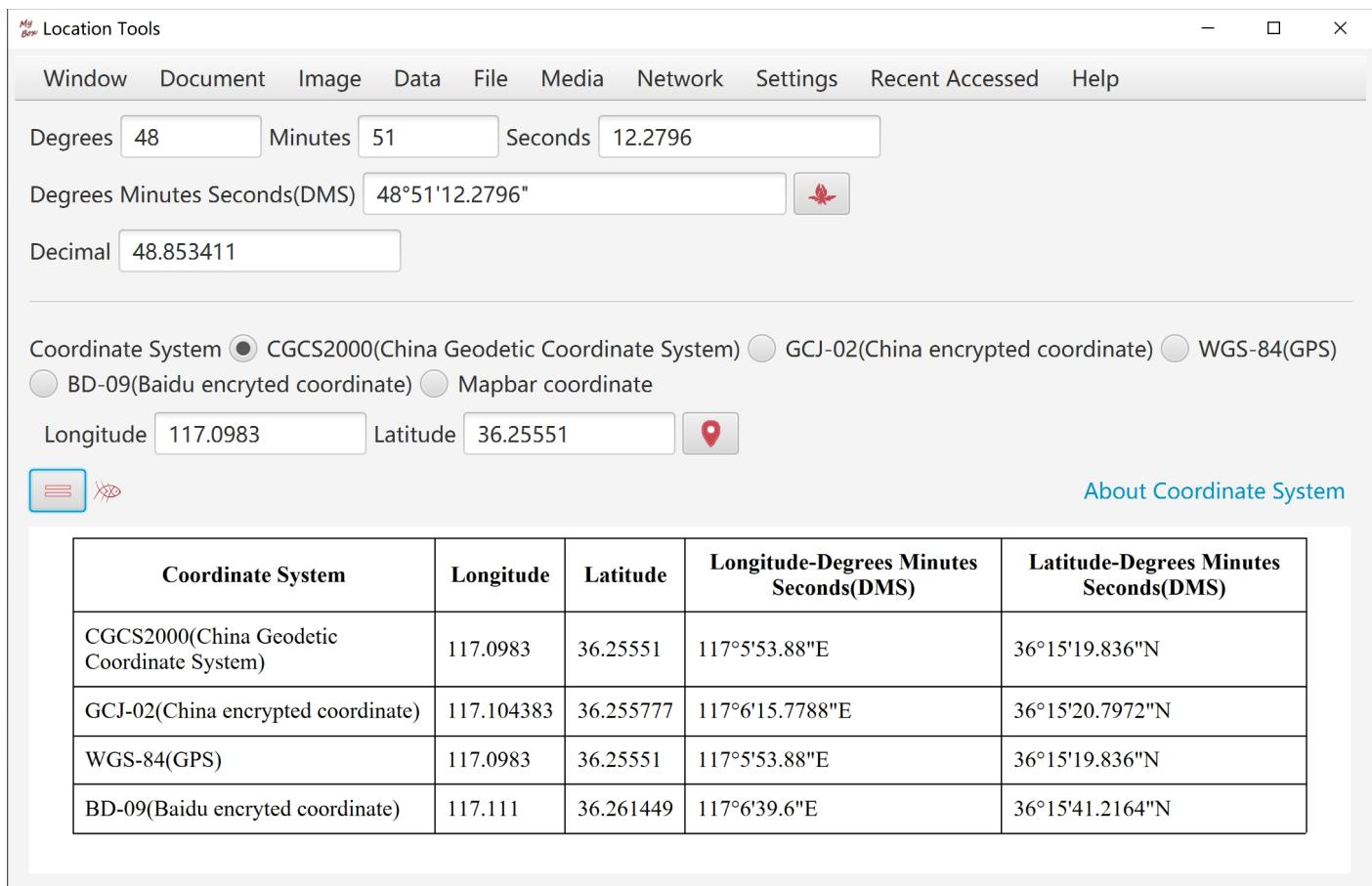


5.6 Location Tools

- Convert coordinate value between decimal and DMS. Valid examples of DMS:

48°51'12.28"
 -77° 3' 43.9308"
 48°51'12.28"N
 2°20'55.68"E
 S 34° 36' 13.4028"
 W 58° 22' 53.7348"
 118 度 48 分 54.152 秒
 -32 度 04 分 10.461 秒
 东经 118 度 48 分 54.152 秒
 北纬 32 度 04 分 10.461 秒
 西经 118 度 48 分 54.152 秒
 南纬 32 度 04 分 10.461 秒

- Convert coordinate values as other coordinate systems.



The screenshot shows the 'Location Tools' window with the following interface elements:

- Menu Bar:** Window, Document, Image, Data, File, Media, Network, Settings, Recent Accessed, Help.
- Coordinate Input:** Degrees (48), Minutes (51), Seconds (12.2796). Below it is a 'Degrees Minutes Seconds(DMS)' input field containing "48°51'12.2796" with a copy icon.
- Coordinate Output:** Decimal (48.853411).
- Coordinate System Selection:** Radio buttons for CGCS2000(China Geodetic Coordinate System) (selected), GCJ-02(China encrypted coordinate), WGS-84(GPS), BD-09(Baidu encrypted coordinate), and Mapbar coordinate.
- Coordinate Fields:** Longitude (117.0983) and Latitude (36.25551) with a location pin icon.
- Buttons:** A menu icon (three horizontal lines), a search icon (magnifying glass), and an 'About Coordinate System' link.
- Table:** A table comparing coordinates across five coordinate systems. The table has columns for Coordinate System, Longitude, Latitude, Longitude-Degrees Minutes Seconds(DMS), and Latitude-Degrees Minutes Seconds(DMS).

Coordinate System	Longitude	Latitude	Longitude-Degrees Minutes Seconds(DMS)	Latitude-Degrees Minutes Seconds(DMS)
CGCS2000(China Geodetic Coordinate System)	117.0983	36.25551	117°5'53.88"E	36°15'19.836"N
GCJ-02(China encrypted coordinate)	117.104383	36.255777	117°6'15.7788"E	36°15'20.7972"N
WGS-84(GPS)	117.0983	36.25551	117°5'53.88"E	36°15'19.836"N
BD-09(Baidu encrypted coordinate)	117.111	36.261449	117°6'39.6"E	36°15'41.2164"N

6 Others

6.1 Create Barcodes

1. Supported 1-d barcodes:
 - Types: Code39, Code128, Codabar, Interleaved2Of5, ITF_14, POSTNET, EAN13, EAN8, EAN_128, UPCA, UPCE, Royal_Mail_Customer_Barcode, USPS_Intelligent_Mail
 - Options about 1-d barcodes: Orientation, width/height, dpi, text location, font size, quiet-zone width, etc.
2. Supported 2-d barcodes:
 - Types: QR_Code, PDF_417, DataMatrix
 - Options about 2-d barcodes: Width/height, margin, error correction level, compact mode, etc.
 - A picture can be shown in center of QR_Code. Its size can be adjusted automatically according to error correction level.
3. Examples of parameters and suggested values.
4. Validate generated barcode at once.

6.2 Decode Barcodes

1. Supported 1-d barcodes: Code39, Code128, Interleaved2Of5, ITF_14, EAN13, EAN8, EAN_128, UPCA, UPCE
2. Supported 2-d barcodes: QR_Code, PDF_417, DataMatrix
3. Display contents of barcodes and its meta data including barcode type and error correction level if any.

6.3 Message Digest

1. Create digest for files or inputted texts.
2. Support MD2, MD5, SHA-1, SHA-224, SHA-256, SHA-384, SHA-512/224, SHA-512/256, SHA3-224, SHA3-256, SHA3-384, SHA3-512.
3. Output: Base64, Hexadecimal, Formatted hexadecimal.

6.4 Encode/Decode Base64

1. Encode file or texts as Base64.
2. Decode Base64 file or Base64 texts.
3. Set charset for texts.
4. Output as file or texts.

6.5 Extract ttf files from ttc file

<End of Document>