



MyBox: Easy Tools Set User Guide – Data Tools

Author: Mara

Version: 6.5.9

Date: 2022-8-31

Contents

1 Introduction.....	6
1.1 Main Interface.....	6
1.2 Resources Addresses.....	7
1.3 Documents.....	8
1.4 Menu of Tools.....	8
2 Two-dimensional Data.....	9
2.1 Edit Data.....	9
2.1.1 Data Objects.....	9
2.1.2 Functions Menu.....	10
2.1.3 Define Columns	11
2.1.4 Define Attributes.....	12
2.1.5 "Table" Edit Mode.....	13
2.1.6 "Text" Edit Mode.....	13
2.1.7 View Html.....	14
2.1.8 View Text.....	14
2.1.9 Load Contents in System Clipboard	15
2.1.10 Import Examples.....	16
2.1.10.1 Statistic Data of China.....	16
2.1.10.2 Data of Regression	16
2.2 Row Expression.....	17
2.3 Row Filter.....	19
2.4 Modify.....	20
2.4.1 Set Values	20
2.4.2 Delete	21
2.4.3 Set styles / Mark Abnormal Values.....	22
2.4.4 Paste Content in System Clipboard.....	23
2.4.5 Paste Content in MyBox Clipboard.....	23
2.5 Trim.....	24
2.5.1 Copy/Filter/Query.....	24
2.5.2 Sort	25
2.5.3 Transpose	26
2.5.4 Normalization	27
2.6 Calculation.....	28
2.6.1 Row Expression.....	28
2.6.2 Descriptive Statistic	29
2.6.3 Group b Equal Values.....	30
2.6.4 Simple Linear Regression	33
2.6.4.1 Regression	33
2.6.4.2 Model.....	34
2.6.4.3 Fitted Chart.....	35
2.6.4.4 Residual Chart.....	36
2.6.5 Simple Linear Regression - Combination	37
2.6.6 Multiple Linear Regression.....	38

2.6.6.1 Regression.....	38
2.6.6.2 Model.....	39
2.6.7 Frequency Distributions	40
2.6.8 Values Percentage	41
2.7 Chart.....	42
2.7.1 XY Chart.....	42
2.7.1.1 Data.....	42
2.7.1.2 Bar Chart.....	43
2.7.1.3 Stacked Bar Chart.....	43
2.7.1.4 Line Chart.....	44
2.7.1.5 Scatter Chart.....	44
2.7.1.6 Bubble Chart.....	45
2.7.1.7 Area Chart.....	46
2.7.1.8 Stacked Area Chart.....	46
2.7.1.9 Parameters of Data in Chart.....	47
2.7.1.10 Layout.....	48
2.7.1.11 Category Axis.....	49
2.7.1.12 Number Axis.....	50
2.7.2 Pie Chart.....	51
2.7.3 Box-and-whisker Chart	52
2.7.4 Self Comparison Bars Chart	54
2.7.5 Comparison Bars Chart	56
2.8 Manage Data.....	58
2.9 Splice Data.....	59
2.10 Data File.....	60
2.10.1 CSV File.....	60
2.10.2 Text File.....	61
2.10.3 Excel File.....	62
2.10.4 Convert/Split Data Files in Batch.....	63
2.10.5 Merge Data Files	64
2.11 Data in System Clipboard.....	65
2.12 Data in MyBox Clipboard.....	66
2.13 Matrix.....	67
2.13.1 Edit and Manage Matrices.....	67
2.13.2 Unary Matrix Calculation.....	68
2.13.3 Binary Matrices Calculation.....	69
2.14 Database Tables.....	70
2.14.1 Manage Database Tables.....	70
2.14.2 Limitations of SQL Identifier.....	71
2.14.3 Database SQL.....	72
2.14.4 SQL Query.....	73
2.15 Export.....	74
2.16 Convert to Database Table.....	75
3 Calculation Tools.....	76
3.1 JShell(Java interactive coding tool)	76
3.2 JEXL(Java Expression Language)	78

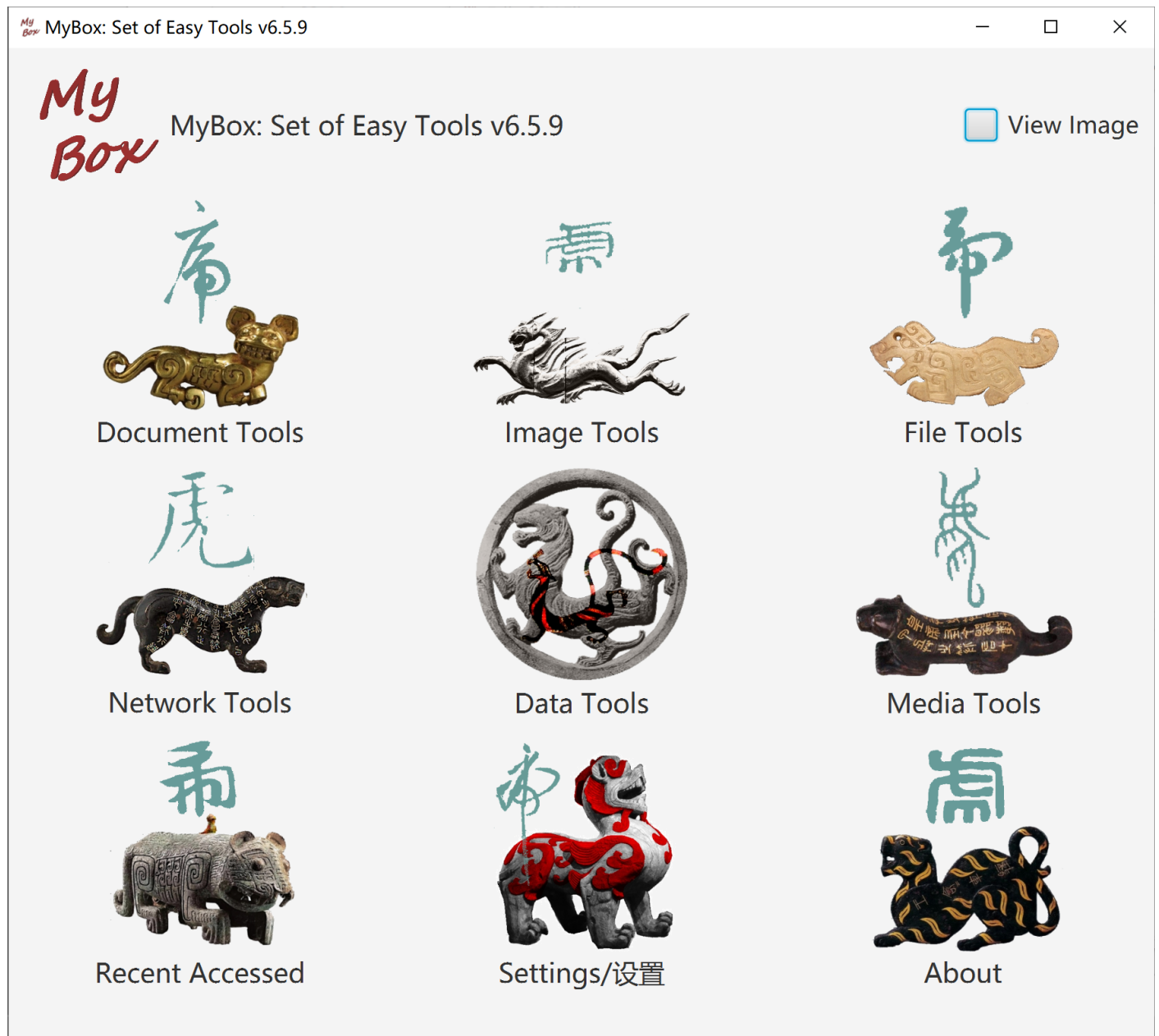
3.3 Javascript	80
4 Data of Location.....	81
4.1 Data Constraints.....	81
4.1.1 Invalid Value.....	81
4.1.2 Coordinate System:.....	81
4.1.3 Coordinate Values.....	81
4.1.4 Time.....	81
4.1.4.1 Date Formats.....	81
4.1.4.2 Era	81
4.1.4.3 Examples	82
4.2 Data Operations.....	82
4.3 Map Data.....	83
4.4 Geography Code.....	85
4.4.1 Data Definition.....	85
4.4.2 Data Constraints.....	85
4.4.3 Edit Data.....	85
4.4.4 Define Condition.....	85
4.4.5 Import Data.....	85
4.4.5.1 Predefined Data.....	85
4.4.5.2 CSV Format.....	86
4.4.5.3 Data from geoname.org:.....	86
4.4.6 Settings	86
4.5 Location in Map.....	88
4.6 Location Data.....	89
4.6.1 Data Definition.....	89
4.6.2 Data constraints.....	89
4.6.3 Define Conditions.....	89
4.6.4 Map Data.....	89
4.6.5 Snapshots.....	90
4.6.6 Import Data.....	90
4.6.6.1 CSV Format.....	90
4.6.6.2 Data from movebank.org.....	90
4.6.6.3 Examples.....	90
4.7 Location Tools.....	92
4.8 Epidemic Reports.....	93
4.8.1 Data Definition.....	93
4.8.2 Data Constraints.....	93
4.8.3 Edit Data.....	93
4.8.4 Import Data.....	94
4.8.4.1 Predefined Data.....	94
4.8.4.2 CSV format:.....	94
4.8.4.3 COVID-19 historical data from Johns Hopkins University(Global) :.....	94
4.8.4.4 COVID-19 daily data from Johns Hopkins University(Global) :.....	94
4.8.4.5 Handle Data of Imported.....	94
4.8.5 Statistics Data.....	95
4.8.6 Define Conditions.....	95

4.8.7 Display Charts.....	95
4.8.8 Settings.....	96
5 Others.....	97
5.1 Create Barcodes.....	97
5.2 Decode Barcodes.....	97
5.3 Message Digest.....	97
5.4 Encode/Decode Base64.....	97
5.5 Extract ttf files from ttc file.....	97

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/MyBox_documents
Mirror Site	https://sourceforge.net/projects/mara-mybox/files/
Cloud Storage	https://pan.baidu.com/s/1fWMRzym_jh075OCX0D8y8A#list/path=%2F

The screenshot shows the GitHub repository page for **Mararsh / MyBox**. The repository is public and has 41 issues, 17 forks, and 79 stars. The main content area displays the repository structure, including folders like `alpha/MyBox`, `docs`, `en`, and `released/MyBox`, as well as files like `.gitignore`, `LICENSE`, and `README.md`. The `README.md` file is open, showing the text "ReadMe in English".

Two red circles highlight specific elements:

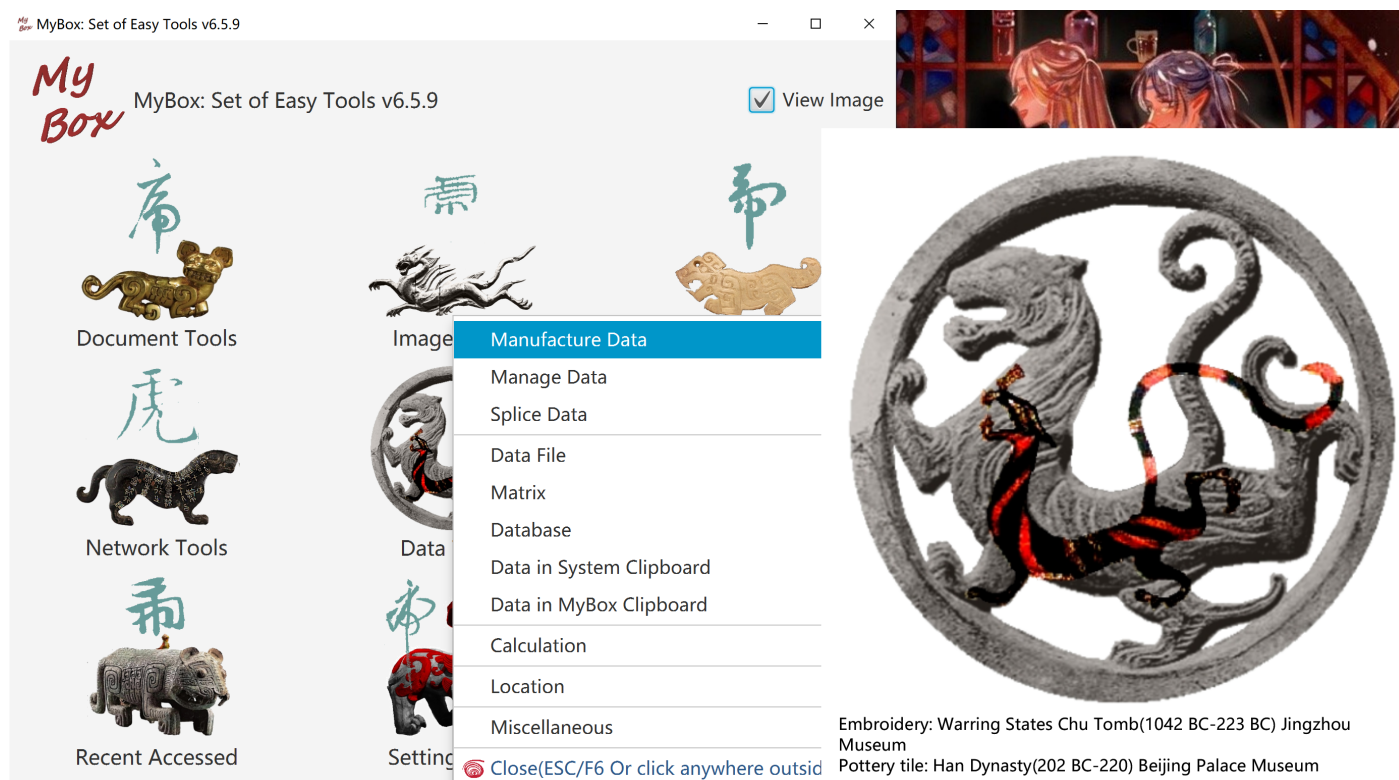
- The **Code** button in the top right corner of the repository view.
- The **Releases** section in the right sidebar, which shows the latest release **v6.5.8** (18 days ago) and a link to **+ 88 releases**.

The URL in the address bar is <https://github.com/Mararsh/MyBox/releases>.

1.3 Documents

Name	Version	Time	English	Chinese
Development Logs	6.5.9	2022-8-31	html	html
Shortcuts	6.5.6	2022-6-11	html	html
Packing Steps	6.3.3	2020-9-27	html	html
Development Guide	2.1	2020-8-27	PDF	PDF
User Guide - Overview	6.5.9	2022-8-31	html PDF odt	html PDF odt
User Guide – Document Tools	6.5.9	2022-8-31	html PDF odt	html PDF odt
User Guide - Image Tools	6.5.9	2022-8-31	html PDF odt	html PDF odt
User Guide - File Tools	6.5.9	2022-8-31	html PDF odt	html PDF odt
User Guide - Network Tools	6.5.9	2022-8-31	html PDF odt	html PDF odt
User Guide - Data Tools	6.5.9	2022-8-31	html PDF odt	html PDF odt
User Guide - Media Tools	6.5.9	2022-8-31	html PDF odt	html PDF odt
User Guide - Development Tools	6.5.9	2022-8-31	html PDF odt	html PDF odt

1.4 Menu of Tools



2 Two-dimensional Data

2.1 Edit Data

2.1.1 Data Objects

- Following objects can be edited in consistent way:
 - Data files, including CSV File, Excel file, texts file.
 - Data of MyBox Clipboard
 - Matrices
 - Database tables.
- Data should be in same width. That is all rows have equal number of columns.
- Data are paginated. When pages number is larger than 1, changes should be saved before run some functions.
- When changed, * is displayed in tab header. And ** is displayed when modifications have not applied.
- 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.
- Click button "Recover" to discard all modifications and load data from file and database.

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

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

Manufacture Data
Manage Data
Splice Data

Data File Edit CSV File
Matrix Convert/Split CSV Files
Database Merge CSV Files
Data in System Clipboard Edit Excel File
Data in MyBox Clipboard Convert/Split Excel Files
Calculation Merge Excel Files
Location Edit Text Data File
Miscellaneous Convert/Split Text Data Files
Merge Text Data Files

Table row	Data row	year						
☐ 1	1	1949						
☐ 2	2	1950						
☐ 3	3	1951						
☐ 4	4	1952						
☐ 5	5	1953						
☐ 6	6	1954						
☐ 7	7	1955						
☐ 8	8	1956						
☐ 9	9	1957	64,655	55,409				
☐ 10	10	1958	65,994	34,195				
☐ 11	11	1959	67,207	34,890	32,317	12,371	54,836	
☐ 12	12	1960	66,207	34,283	31,924	13,073	53,134	
☐ 13	13	1961	65,859	33,880	31,979	12,707	53,152	

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

2.1.2 Functions Menu

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

ColumnsAttributesViewEdit

row

Save

Recover

Refresh

Open

Create data

Load content in system clipboard.

Export

Convert to database table

Data

Modify

Trim

Calculation

Charts

Examples

☒ Pop when mouse hovering

Close(ESC/F6 Or click anywhere outside the object)

1956

62,828

32,338

30,292

9,183

53,843

1957

64,653

33,469

31,184

9,949

54,704

1958

65,994

34,195

31,799

10,721

55,273

Hover or click to pop menu

2.1.3 Define Columns

1. Column names should not be null nor duplicated.
2. Data types are used to validate data values:
 - Invalid value is rejected when edit data.
 - Type is ignored when read or calculate data.
 - Data type affects sorting results.
3. Click button "OK" to apply it modifications to "Table".
4. Click button "Cancel" to discard its modifications and pick data from "Table".
5. Can rename all columns with sequence numbers.
6. Can set random colors.

Columns Attributes View Edit									
☐ Table row	Index	Column name	Type	Color	Width	Editable	Not null	Format	Length +
☐ 1	0	year_	String		100	☒	☐	☒	32,672
☐ 2	1	population at year-end(ten thousand)	Double		100	☒	☐	☒	32,672
☐ 3	2	male(ten thousand)	Double		100	☒	☐	☒	32,672
☐ 4	3	female(ten thousand)	Double		100	☒	☐	☒	32,672
☒ 5	4	urban(ten thousand)	Double		100	☒	☐	☒	32,672
☐ 6	5	rural(ten thousand)	Double		100	☒	☐	☒	32,672

2.1.4 Define Attributes

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

The screenshot shows the 'Attributes' tab in the MyBox Data Tools v6.5.9 interface. The tab is highlighted with a blue circle. The interface displays the following information:

- CSV - 786 - d:\tmp\mybox-data-6.5.9\data\ChinaPopulation_en.csv**
- ID** 786
- Data type** CSV
- Data name** ChinaPopulation_en.csv
- Decimal scale** 2
- Maximum value of random** 1000
- Update time** 2022-08-30 17:56:32

Below these fields, a text area displays file details:

- File: d:\tmp\mybox-data-6.5.9\data\ChinaPopulation_en.csv
- File size: 2.719 KB
- File modify time: 2022-08-30 17:56:32
- Charset: US-ASCII
- Delimiter: ,
- First line defines the columns' names: Yes
- Lines number in file: 73
- Columns number: 6
- Current page: 1 / 2

At the bottom, a status bar shows: Selected: 0 Rows: 60/73 Page size: 60 Page: 1 / 2. Navigation icons are visible on the right side of the status bar.

2.1.5 "Table" Edit Mode

"Table" is the master edit mode:

1. Its modifications are applied to other panes automatically.
2. It is the final data to save.

Columns* Attributes View Edit					
year_	population at year-end...	male(ten thousan...	female(ten thous...	urban(ten +	
1949	54,167	28,145	26,022	5,765	
1950	55,196	28,669	26,527	6,169	
1951	56,300	29,231	27,069	6,632	
1952	57,482	29,833	27649	7163	
1953	58,796	30,468	28,328	7,826	
1954	60,266	31,242	29,024	8,249	

2.1.6 "Text" Edit Mode

"Text" 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.

Columns* Attributes View Edit

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

1949|54167.0|28145.0|26022.0|5765.0|48402.0
1950|55196.0|28669.0|26527.0|6169.0|49027.0
1951|56300.0|29231.0|27069.0|6632.0|49668.0
1952|57482.0|29833.0|27649.0|7163.0|50319.0
1953|58796.0|30468.0|28328.0|7826.0|50970.0
1954|60266.0|31242.0|29024.0|8249.0|52017.0
1955|61465.0|31809.0|29656.0|8285.0|53180.0
1956|62828.0|32536.0|30292.0|9185.0|53643.0
1957|64653.0|33469.0|31184.0|9949.0|54704.0
1958|65994.0|34195.0|31799.0|10721.0|55273.0
1959|67207.0|34890.0|32317.0|12371.0|54836.0
1960|66207.0|34283.0|31924.0|13073.0|53134.0
1961|65859.0|33880.0|31979.0|12707.0|53152.0
1962|67296.0|34517.0|32778.0|11659.0|55636.0
1963|69172.0|35533.0|33639.0|11646.0|57526.0

MyBox: Set of Easy Tools

Text delimiter

☐ String ☐ Tab ☐ Blank ☐ Blank characters ☐ 4 blanks ☐ 8 blanks

☐ , ☐ . ☐ # ☐ - ☐ + ☐ : ☐ ; ☐ @ ☐ & ☐ % ☐ ! ☐ " ☐ ' ☐ ? ☐ * ☐ \ ☐ / ☐ _ ☐ = ☐ < ☐ >

2.1.7 View Html

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

Columns* Attributes View Edit					
☐ Form	☐ Title	☒ Column name	☐ Row number		
year_	population at year-end(ten thousand)	male(ten thousand)	female(ten thousand)	urban(ten thousand)	rural(ten thousand)
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

2.1.8 View Text

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

Columns* Attributes View Edit					
☐ Form	☐ Title	☒ Column name	☒ Row number		

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

Row1#1949#54167.0#28145.0#26022.0#5765.0#48402.0

Row2#1950#55196.0#28669.0#26527.0#6169.0#49027.0

Row3#1951#56300.0#29231.0#27069.0#6632.0#49668.0

Row4#1952#57482.0#29833.0#27649.0#7163.0#50319.0

Row5#1953#58796.0#30468.0#28328.0#7826.0#50970.0

Row6#1954#60266.0#31242.0#29024.0#8249.0#52017.0

Row7#1955#61465.0#31809.0#29656.0#8285.0#53180.0

Row8#1956#62828.0#32536.0#30292.0#9185.0#53643.0

Row9#1957#64653.0#33469.0#31184.0#9949.0#54704.0

Row10#1958#65994.0#34195.0#31799.0#10721.0#55273.0

Row11#1959#67207.0#34890.0#32317.0#12371.0#54836.0

Row12#1960#66207.0#34283.0#31924.0#13073.0#53134.0

Row13#1961#65859.0#33880.0#31979.0#12707.0#53152.0

Row14#1962#67599.0#35117.0#33478.0#14168.0#55431.0

Row15#1963#69399.0#36465.0#34886.0#15213.0#57186.0

Row16#1964#71299.0#37813.0#36394.0#16367.0#59432.0

Row17#1965#73299.0#39213.0#37902.0#17519.0#61780.0

Row18#1966#75299.0#40613.0#39410.0#18671.0#64128.0

Row19#1967#77299.0#42013.0#40918.0#19823.0#66476.0

Row20#1968#79299.0#43413.0#42426.0#20975.0#68824.0

Row21#1969#81299.0#44813.0#43934.0#22127.0#71172.0

Row22#1970#83299.0#46213.0#44942.0#23279.0#73520.0

Row23#1971#85299.0#47613.0#45950.0#24431.0#75868.0

Row24#1972#87299.0#49013.0#46958.0#25583.0#78216.0

Row25#1973#89299.0#50413.0#47966.0#26735.0#80564.0

Row26#1974#91299.0#51813.0#48974.0#27887.0#82912.0

Row27#1975#93299.0#53213.0#50482.0#29039.0#85260.0

Row28#1976#95299.0#54613.0#51490.0#30191.0#87608.0

Row29#1977#97299.0#56013.0#52498.0#31343.0#89956.0

Row30#1978#99299.0#57413.0#53506.0#32495.0#92304.0

Row31#1979#101299.0#58813.0#54514.0#33647.0#94652.0

Row32#1980#103299.0#60213.0#55522.0#34799.0#97000.0

Row33#1981#105299.0#61613.0#56530.0#35951.0#99348.0

Row34#1982#107299.0#63013.0#57538.0#37103.0#101696.0

Row35#1983#109299.0#64413.0#58546.0#38255.0#104044.0

Row36#1984#111299.0#65813.0#59554.0#39407.0#106392.0

Row37#1985#113299.0#67213.0#60562.0#40559.0#108740.0

Row38#1986#115299.0#68613.0#61570.0#41711.0#111088.0

Row39#1987#117299.0#70013.0#62578.0#42863.0#113436.0

Row40#1988#119299.0#71413.0#63586.0#44015.0#115784.0

Row41#1989#121299.0#72813.0#64594.0#45167.0#118132.0

Row42#1990#123299.0#74213.0#65602.0#46319.0#120480.0

Row43#1991#125299.0#75613.0#66610.0#47471.0#122828.0

Row44#1992#127299.0#77013.0#67618.0#48623.0#125176.0

Row45#1993#129299.0#78413.0#68626.0#49775.0#127524.0

Row46#1994#131299.0#79813.0#69634.0#50927.0#129872.0

Row47#1995#133299.0#81213.0#70642.0#52079.0#132220.0

Row48#1996#135299.0#82613.0#71650.0#53231.0#134568.0

Row49#1997#137299.0#84013.0#72658.0#54383.0#136916.0

Row50#1998#139299.0#85413.0#73666.0#55535.0#139264.0

Row51#1999#141299.0#86813.0#74674.0#56687.0#141612.0

Row52#2000#143299.0#88213.0#75682.0#57839.0#143960.0

Row53#2001#145299.0#89613.0#76690.0#58991.0#146308.0

Row54#2002#147299.0#91013.0#77698.0#60143.0#148656.0

Row55#2003#149299.0#92413.0#78706.0#61295.0#151004.0

Row56#2004#151299.0#93813.0#79714.0#62447.0#153352.0

Row57#2005#153299.0#95213.0#80722.0#63599.0#155700.0

Row58#2006#155299.0#96613.0#81730.0#64651.0#158048.0

Row59#2007#157299.0#98013.0#82738.0#65803.0#160396.0

Row60#2008#159299.0#99413.0#83746.0#66955.0#162744.0

Row61#2009#161299.0#100813.0#84754.0#68107.0#165092.0

Row62#2010#163299.0#102213.0#85762.0#69259.0#167440.0

Row63#2011#165299.0#103613.0#86770.0#70411.0#169788.0

Row64#2012#167299.0#105013.0#87778.0#71563.0#172136.0

Row65#2013#169299.0#106413.0#88786.0#72715.0#174484.0

Row66#2014#171299.0#107813.0#89794.0#73867.0#176832.0

Row67#2015#173299.0#109213.0#90802.0#75019.0#179180.0

Row68#2016#175299.0#110613.0#91810.0#76171.0#181528.0

Row69#2017#177299.0#112013.0#92818.0#77323.0#183876.0

Row70#2018#179299.0#113413.0#93826.0#78475.0#186224.0

Row71#2019#181299.0#114813.0#94834.0#79627.0#188572.0

Row72#2020#183299.0#116213.0#95842.0#80779.0#190920.0

Row73#2021#185299.0#117613.0#96850.0#81931.0#193268.0

Row74#2022#187299.0#119013.0#97858.0#83083.0#195616.0

Row75#2023#189299.0#120413.0#98866.0#84235.0#197964.0

Row76#2024#191299.0#121813.0#99874.0#85387.0#200312.0

Row77#2025#193299.0#123213.0#100882.0#86539.0#202660.0

Row78#2026#195299.0#124613.0#101890.0#87691.0#205008.0

Row79#2027#197299.0#126013.0#102898.0#88843.0#207356.0

Row80#2028#199299.0#127413.0#103906.0#89995.0#209704.0

Row81#2029#201299.0#128813.0#104914.0#91147.0#212052.0

Row82#2030#203299.0#130213.0#105922.0#92299.0#214400.0

Row83#2031#205299.0#131613.0#106930.0#93451.0#216748.0

Row84#2032#207299.0#133013.0#107938.0#94603.0#219096.0

Row85#2033#209299.0#134413.0#108946.0#95755.0#221444.0

Row86#2034#211299.0#135813.0#109954.0#96907.0#223792.0

Row87#2035#213299.0#137213.0#110962.0#98059.0#226140.0

Row88#2036#215299.0#138613.0#111970.0#99211.0#228488.0

Row89#2037#217299.0#140013.0#112978.0#100363.0#230836.0

Row90#2038#219299.0#141413.0#113986.0#101515.0#233184.0

Row91#2039#221299.0#142813.0#114994.0#102667.0#235532.0

Row92#2040#223299.0#144213.0#115999.0#103819.0#237880.0

Row93#2041#225299.0#145613.0#117007.0#104971.0#240228.0

Row94#2042#227299.0#147013.0#118015.0#106123.0#242576.0

Row95#2043#229299.0#148413.0#119023.0#107275.0#244924.0

Row96#2044#231299.0#149813.0#120031.0#108427.0#247272.0

Row97#2045#233299.0#151213.0#121039.0#109579.0#249620.0

Row98#2046#235299.0#152613.0#122047.0#110731.0#251968.0

Row99#2047#237299.0#154013.0#123055.0#111883.0#254316.0

Row100#2048#239299.0#155413.0#124063.0#113035.0#256664.0

Row101#2049#241299.0#156813.0#125071.0#114187.0#259012.0

Row102#2050#243299.0#158213.0#126079.0#115339.0#261360.0

Row103#2051#245299.0#159613.0#127087.0#116491.0#263708.0

Row104#2052#247299.0#161013.0#128095.0#117643.0#266056.0

Row105#2053#249299.0#162413.0#129103.0#118795.0#268404.0

Row106#2054#251299.0#163813.0#130111.0#119947.0#270752.0

Row107#2055#253299.0#165213.0#131119.0#121099.0#273100.0

Row108#2056#255299.0#166613.0#132127.0#122251.0#275448.0

Row109#2057#257299.0#168013.0#133135.0#123403.0#277796.0

Row110#2058#259299.0#169413.0#134143.0#124555.0#280144.0

Row111#2059#261299.0#170813.0#135151.0#125707.0#282492.0

Row112#2060#263299.0#172213.0#136159.0#126859.0#284840.0

Row113#2061#265299.0#173613.0#137167.0#128011.0#287188.0

Row114#2062#267299.0#175013.0#138175.0#129163.0#289536.0

Row115#2063#269299.0#176413.0#139183.0#130315.0#291884.0

Row116#2064#271299.0#177813.0#140191.0#131467.0#294232.0

Row117#2065#273299.0#179213.0#141199.0#132619.0#296580.0

Row118#2066#275299.0#180613.0#142207.0#133771.0#298928.0

Row119#2067#277299.0#182013.0#143215.0#134923.0#301276.0

Row120#2068#279299.0#183413.0#144223.0#136075.0#303624.0

Row121#2069#281299.0#184813.0#145231.0#137227.0#305972.0

Row122#2070#283299.0#186213.0#146239.0#138379.0#308320.0

Row123#2071#285299.0#187613.0#147247.0#139531.0#310668.0

Row124#2072#287299.0#189013.0#148255.0#140683.0#313016.0

Row125#2073#289299.0#190413.0#149263.0#141835.0#315364.0

Row126#2074#291299.0#191813.0#150271.0#142987.0#317712.0

Row127#2075#293299.0#193213.0#151279.0#144139.0#320060.0

Row128#2076#295299.0#194613.0#152287.0#145291.0#322408.0

Row129#2077#297299.0#196013.0#153295.0#146443.0#324756.0

Row130#2078#299299.0#197413.0#154303.0#147595.0#327104.0

Row131#2079#299299.0#197413.0#154303.0#147595.0#327104.0

Row132#2080#299299.0#197413.0#154303.0#147595.0#327104.0

Row133#2081#299299.0#197413.0#154303.0#147595.0#327104.0

Row134#2082#299299.0#197413.0#154303.0#147595.0#327104.0

Row135#2083#299299.0#197413.0#154303.0#147595.0#327104.0

Row136#2084#299299.0#197413.0#154303.0#147595.0#327104.0

Row137#2085#299299.0#197413.0#154303.0#147595.0#327104.0

Row138#2086#299299.0#197413.0#154303.0#147595.0#327104.0

Row139#2087#299299.0#197413.0#154303.0#147595.0#327104.0

Row140#2088#299299.0#197413.0#154303.0#147595.0#327104.0

Row141#2089#299299.0#197413.0#154303.0#147595.0#327104.0

Row142#2090#299299.0#197413.0#154303.0#147595.0#327104.0

Row143#2091#299299.0#197413.0#154303.0#147595.0#327104.0

Row144#2092#299299.0#197413.0#154303.0#147595.0#327104.0

Row145#2093#299299.0#197413.0#154303.0#147595.0#327104.0

Row146#2094#299299.0#197413.0#154303.0#147595.0#327104.0

Row147#2095#299299.0#197413.0#154303.0#147595.0#327104.0

Row148#2096#299299.0#197413.0#154303.0#147595.0#327104.0

Row149#2097#299299.0#197413.0#154303.0#147595.0#327104.0

Row150#2098#299299.0#197413.0#154303.0#147595.0#327104.0

Row151#2099#299299.0#197413.0#154303.0#147595.0#327104.0

Row152#2100#299299.0#197413.0#154303.0#147595.0#327104.0

Row153#2101#299299.0#197413.0#154303.0#147595.0#327104.0

Row154#2102#299299.0#197413.0#154303.0#147595.0#327104.0

Row155#2103#299299.0#197413.0#154303.0#147595.0#327104.0

Row156#2104#299299.0#197413.0#154303.0#147595.0#327104.0

Row157#2105#299299.0#197413.0#154303.0#147595.0#327104.0

Row158#2106#299299.0#197413.0#154303.0#147595.0#327104.0

Row159#2107#299299.0#197413.0#154303.0#147595.0#327104.0

Row160#2108#299299.0#197413.0#154303.0#147595.0#327104.0

Row161#2109#299299.0#197413.0#154303.0#147595.0#327104.0

Row162#2110#299299.0#197413.0#154303.0#147595.0#327104.0

Row163#2111#299299.0#197413.0#154303.0#147595.0#327104.0

Row164#2112#299299.0#197413.0#154303.0#147595.0#327104.0

Row165#2113#299299.0#197413.0#154303.0#147595.0#327104.0

Row166#2114#299299.0#197413.0#154303.0#147595.0#327104.0

Row167#2115#299299.0#197413.0#154303.0#147595.0#327104.0

Row168#2116#299299.0#197413.0#154303.0#147595.0#327104.0

Row169#2117#299299.0#197413.0#154303.0#147595.0#327104.0

Row170#2118#299299.0#197413.0#154303.0#147595.0#327104.0

Row171#2119#299299.0#197413.0#154303.0#147595.0#327104.0

Row172#2120#299299.0#197413.0#154303.0#147595.0#327104.0

Row173#2121#299299.0#197413.0#154303.0#147595.0#327104.0

Row174#2122#299299.0#197413.0#154303.0#147595.0#327104.0

Row175#2123#299299.0#197413.0#154303.0#147595.0#327104.0

Row176#2124#299299.0#197413.0#154303.0#147595.0#327104.0

Row177#2125#299299.0#197413.0#154303.0#147595.0#327104.0

Row178#2126#299299.0#197413.0#154303.0#147595.0#327104.0

Row179#2127#299299.0#197413.0#154303.0#147595.0#327104.0

Row180#2128#299299.0#197413.0#154303.0#147595.0#327104.0

Row181#2129#299299.0#197413.0#154303.0#147595.0#327104.0

Row182#2130#299299.0#197413.0#154303.0#147595.0#327104.0

Row183#2131#299299.0#197413.0#154303.0#147595.0#327104.0

Row184#2132#299299.0#197413.0#154303.0#147595.0#327104.0

Row185#2133#299299.0#197413.0#154303.0#147595.0#327104.0

Row186#2134#299299.0#197413.0#154303.0#147595.0#327104.0

Row187#2135#299299.0#197413.0#154303.0#147595.0#327104.0

Row188#2136#299299.0#197413.0#154303.0#147595.0#327104.0

Row189#2137#299299.0#197413.0#154303.0#147595.0#327104.0

Row190#2138#299299.0#197413.0#154303.0#147595.0#327104.0

Row191#2139#299299.0#197413.0#154303.0#147595.0#327104.0

Row192#2140#299299.0#197413.0#154303.0#147595.0#327104.0

Row193#2141#299299.0#197413.0#154303.0#147595.0#327104.0

Row194#2142#299299.0#197413.0#154303.0#147595.0#327104.0

Row195#2143#299299.0#197413.0#154303.0#147595.0#327104.0

Row196#21

2.1.9 Load Contents in System Clipboard

1. Read and parse contents in system clipboard.
2. Guess delimiter automatically.
3. Delimiter can be chosen from special characters or inputted regular expression.
4. First row can be set as column names.

MyBox Load content in system clipboard.

☒ First line defines the columns' names

Delimiter

Source row number	year_	population at year-end(ten thousand)	male(ten thousand)	female(ten thousand)
Row1	1949	54167	28145	26022
Row2	1950	55196	28669	26527
Row3	1951	56300	29231	27069
Row4	1952	57482	29833	27649
Row5	1953	58796	30468	28328
Row6	1954	60266	31242	29024
Row7	1955	61465	31809	29656
Row8	1956	62828	32536	30292

year_,population at year-end(ten thousand),r
 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

2.1.10 Import Examples

2.1.10.1 Statistic Data of China

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

Window Document Image Data File Media Network Settings Recent Accessed






 **Statistic data of China**

 **Regression**

☒  Pop when mouse hovering

 Close(ESC/F6 Or click anywhere outside the object)

☐ 2	2	1950	55,196	28,669
☐ 3	3	1951	56,300	29,231
☐ 4	4	1952	57,482	29,833
☐ 5	5	1953	58,796	30,468
☐ 6	6	1954	60,266	31,242
☐ 7	7	1955	61,465	31,809
☐ 8	8	1956	62,828	32,536
☐ 9	9	1957	64,653	33,469
☐ 10	10	1958	65,994	34,195
☐ 11	11	1959	67,207	34,890
☐ 12	12	1960	66,207	34,283

- 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**
- Health personnel of China
- Marriage of China
- Sport world champions of China
- Crimes filed by China police
- Crimes filed by China procuratorate
- China National Bureau of Statistics

2.1.10.2 Data of Regression






 **Statistic data of China**

 **Regression**

☒  Pop when mouse hovering

 Close(ESC/F6 Or click anywhere outside the object)

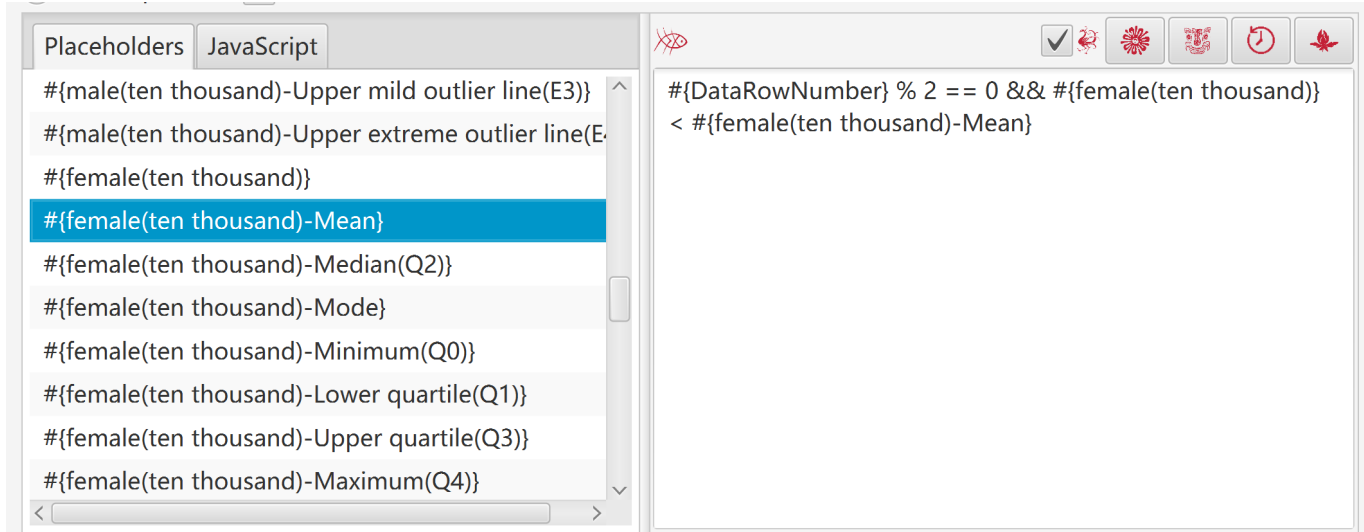
☐ 2	2	1950	55,196	28,669
☐ 3	3	1951	56,300	29,231

- Income and happiness
- Years experience and salary
- Iris species
- Boston housing prices

2.2 Row Expression

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

1. If the script is blank, then return empty string.
2. Edit the script:
 - It can include any valid JavaScript elements.
 - It should return a value finally.
 - It can include following placeholders:
 - `#{TableRowNumber}`
 - `#{DataRowNumber}`
 - `#{<column_name>}`
 - `#{<column_name>- }`
3. When MyBox evaluates the expression:
 - Placeholders are replaced with actual values of each data row .
 - `'#{xxx}'` is handled as string while `#{xxx}` is handled as number.
 - When handles all pages, script fails when it includes “`#{TableRowNumber}`” .
4. Hover or click button “Examples” to paste example codes.
5. Hover or click button “Histories” to paste codes in histories.



6. 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}) + Math.PI * Math.sqrt({v2})</code>	calculation
<code>'{v1}'.replace(/hello/ig, 'Hello')</code>	replace all “hello”(case-insensitive) as “Hello” in column “v1”
<code>'{v1}'.toLowerCase()</code>	lower case of value of column “v1”
<code>'{v1}'.split(',')</code>	split value of column “v1” by comma
<code>{v1} - {v1-Mean}</code>	difference between value of column “v1” and mean of column “v1”

2.3 Row Filter

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

1. It should return boolean value("true" or "false") finally.
2. Can set maximum rows.

Maximum rows to take(Empty/zero/negative to unlimited) 5

Select to paste

Placeholders JavaScript

Take source data when the expression is blank or ☒ true ☐ Others

`#{TableRowNumber} % 2 == 0 && #{year_} > 1995`

Expressions

- Numeric operations
- Strings operations
- Array operations
- Boolean operations
 - string include
 - string match
 - array include
 - is number
 - and/or/not

3. Exmaples:

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}'.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”

2.4 Modify

2.4.1 Set Values

1. Select 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.
2. Set values of selected data as following:
 - Constant: 0, 1, blank, or inputted value
 - Random, random of non-negative
 - Add prefix, append suffix, set decimal scale, or add sequence numbers
 - 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.
 - Row expression
3. If handle all pages of data file, then auto-backup before set values.

Set values - CSV - 786 - d:\tmp\mybox-data-6.5.9\data\ChinaPopulation_en.csv

Set values

Data Row filter Columns Values

☐ 0 ☐ 1 ☐ Blank ☐ Random ☐ Random(non-negative)

☐ Value ☐ Add prefix ☐ Append suffix

☐ Decimal scale To invalid number ☒ Skip ☐ Zero ☐ Blank

☒ Add sequence number Start ☒ Fill zero(like 001,0001) Digit (Empty to handle automatically)

☐ Gaussian distribution ☐ Identify matrix ☐ Upper triangle ☐ Lower triangle

☐ Row expression ☒ Continue when error

Placeholders JavaScript

#{female(ten thousand)-Mean}
 #{female(ten thousand)-Median(Q2)}
 #{female(ten thousand)-Mode}
 #{female(ten thousand)-Minimum(Q0)}
 #{female(ten thousand)-Lower quartile(Q1)}

#{DataRowNumber} % 2 == 0 && #{female(ten thousand)} < #{female(ten thousand)-Mean}

2.4.2 Delete

1. Select data:
 - Rows can be: current page, selected rows, or all pages.
 - Set row filter.
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

Table row	Data row	n(te... rural(ten ...
☐ 1	1	
☐ 2	2	
☐ 3	3	
☐ 4	4	
☐ 5	5	
☐ 6	6	
☐ 7	7	
☐ 8	8	
☐ 9	9	
☐ 10	10	
☐ 11	11	
☐ 12	12	
☐ 13	13	
☐ 14	14	
☐ 15	15	
☐ 16	16	

Select some to handle, or select none to handle all in table:

- 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.

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

2.4.3 Set styles / Mark Abnormal Values

1. Add/Edit/Delete styles.
2. Define conditions to determine which data cells to apply the style:
 - Range of data rows
 - Column names.
 - 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 sepcial rows is the expression composed of column values.

3. Define style values:
 - Font color, font size, background color, bold, etc.
 - More values in format of JavaFx CSS.
4. Define title and sequence number of the style.
5. Set whether the style marks abnormal values.
6. All styles are applied to the data one by one in order of their sequence number.

Set Styles / Mark Abnormal Values - CSV - 786 - d:\tmp\mybox-data-6.5.9\data\ChinaPopulation_en.csv

Set Styles / Mark Abnormal Values

Se...	Abnormal...	Title	+
1.0	✓	Female A	
2.0	✓	Male A	
3.0		Female B	
4.0		Male B	

ID: 5

Base Data Filter Style

Font color ☐ Default ☒ Set

Background color ☐ Default ☒ Set

Font size ☐ Bold

More

Page 1 / 1

Page size 50

Rows: 4/4 Selected: 1

[JavaFX CSS Reference Guide](#)

2.4.4 Paste Content in System Clipboard

MyBox Paste content in System Clipboard

☒ First line defines the columns' names

Location to paste
Table row 1
Column year_

☐ Replace
☐ Insert above
☐ Append below

Delimiter

Source row number	year_	population at year-end(ten thousand)	male(ten thousand)	female(ten thousand)
Row1	1951	56300	29231	27069
Row2	1952	57482	29833	27649
Row3	1953	58796	30468	28328
Row4	1954	60266	31242	29024
Row5	1955	61465	31809	29656

year_,population at year-end(ten thousand)
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

2.4.5 Paste Content in MyBox Clipboard

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

☒ First line defines the columns' names

Location to paste
Table row 1
Column year_

☐ Replace
☐ Insert above
☐ Append below

Table row ID Type Name

Table row	ID	Type	Name
☒ 1	1089	MyBox ...	ChinaPopu
☐ 2	1086	MyBox ...	ChinaPopu
☐ 3	1085	MyBox ...	ChinaPopu
☐ 4	1083	MyBox ...	ChinaPopu
☐ 5	1082	MyBox ...	ChinaPopu
☐ 6	1063	MyBox ...	a
☐ 7	6	MyBox ...	b

Page 1 /1
Page size 50
Rows: 7/7 Selected: 1

Data Row filter

Rows ☒ Selected ☐ Current page ☐ All pages

Columns

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

2.5 Trim

2.5.1 Copy/Filter/Query

1. Select 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.
2. Target can be following: new csv/excel/text file, matrix, system clipboard, MyBox clipboard, database table.
3. When rows are current page or selected ones, target can be defined location in table to insert/append/replace.

MyBox Copy/Filter/Query - CSV - 786 - d:\tmp\mybox-data-6.5.9\data\ChinaPopulation_en.csv

Copy/Filter/Query

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

Columns   No selection means all columns

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

☐ Row number ☒ Column name

Target

External ☐       

Name

In table ☐ Insert above ☐ Append below ☐ Replace

2.5.2 Sort

1. Select 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.
2. Calculate:
 - Select the column to sort and whether descending.
 - Data type of column affects sorting results.
 - Maximum rows number of results can be set.
3. Target can be following: new csv/excel/text file, matrix, system clipboard, MyBox clipboard, database table.
4. When rows are current page or selected ones, target can be defined location in table to insert/append/replace.

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

Sort

Data Row filter Options

Order by(Column type affects sorting results)

Table row	Column	
☐ 1	year_-Descending	
☐ 2	year_-Ascending	
☒ 3	population at year-end(ten thousand)-Descending	
☐ 4	population at year-end(ten thousand)-Ascending	
☐ 5	male(ten thousand)-Descending	

Columns  ☐ No selection means all columns

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

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

☐ Row number ☒ Column name

Target

External ☒ CSV ☐ XLS ☐ TXT ☐  ☐  ☐ 

Name

In table ☐ Insert above ☐ Append below ☐ Replace

2.5.3 Transpose

1. Select 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.
2. Calculate: Options to set first column as column names.
3. Target can be following: new csv/excel/text file, matrix, system clipboard, MyBox clipboard, database table.
4. When rows are current page or selected ones, target can be defined location in table to insert/append/replace.

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

Transpose

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

Columns   No selection means all columns

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

☐ urban(ten thousand) ☐ rural(ten thousand)

☐ Row number ☒ Column name ☒ First column as column names

Target

External ☐        

In table ☐ Insert above ☒ Append below ☐ Replace

Table row

Column

2.5.4 Normalization

1. Select 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.
2. Calculate:
 - According to: Columns/rows/all.
 - Algorithms: MinMax(Range can be set), sum(L1), ZScore(L2).
 - To non-numeric, skip or count as zero.
3. Target can be following: new csv/excel/text file, matrix, system clipboard, MyBox clipboard, database table.
4. When rows are current page or selected ones, target can be defined location in table to insert/append/replace.

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

Normalize

Data
Row filter
Options

Columns   No selection means all columns

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

According to ☒ Columns ☐ Rows ☐ All

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

Range ,

☐ Row number
 ☒ Column name

Decimal scale ▼

To invalid number ☒ Skip ☐ Count as zero

Target

External          

Name

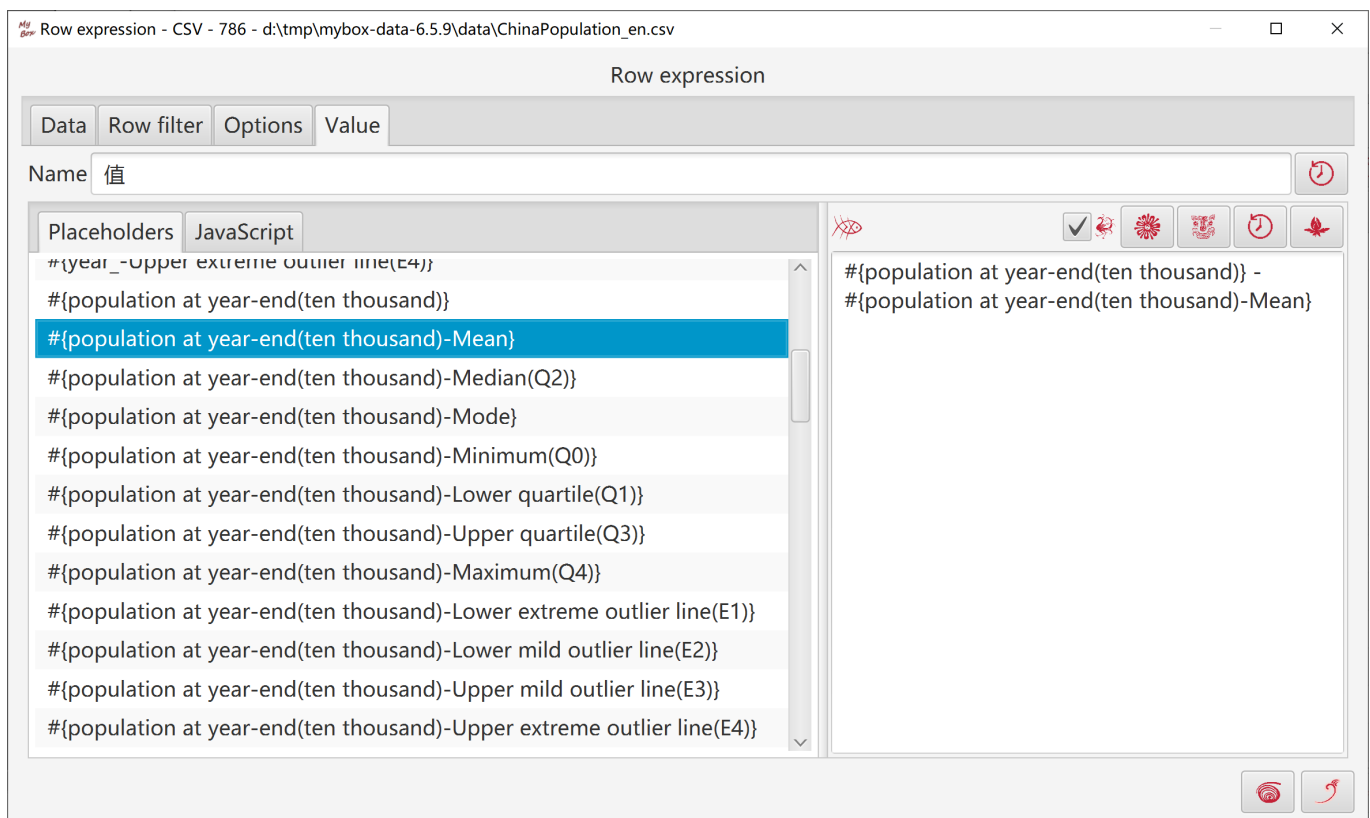
In table ☐ Insert above ☐ Append below ☐ Replace




2.6 Calculation

2.6.1 Row Expression

1. Select 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.
2. Calculate:
 - Input name of values.
 - Input row expression.
3. Target can be following: new csv/excel/text file, matrix, system clipboard, MyBox clipboard, database table.
4. When rows are current page or selected ones, target can be defined location in table to insert/append/replace.



2.6.2 Descriptive Statistic

1. Select 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.

2. Calculate:

- Following values can be generated:
count, sum, 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.
- Set decimal scale.
- To non-numeric, skip or count as zero.

3. Target can be following: new csv/excel/text file, matrix, system clipboard, MyBox clipboard, database table, or defined location in table to insert/append/replace.

Descriptive statistics - CSV - 1103 - d:\tmp\mybox-data-6.5.9\data\ChinaPopulation_en.csv

Descriptive statistics

Data Row filter Options

Columns   No selection means all columns

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

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

To invalid number ☒ Skip ☐ Count as zero

Target

External ☒       

Name

In table ☐ Insert above ☐ Append below ☐ Replace

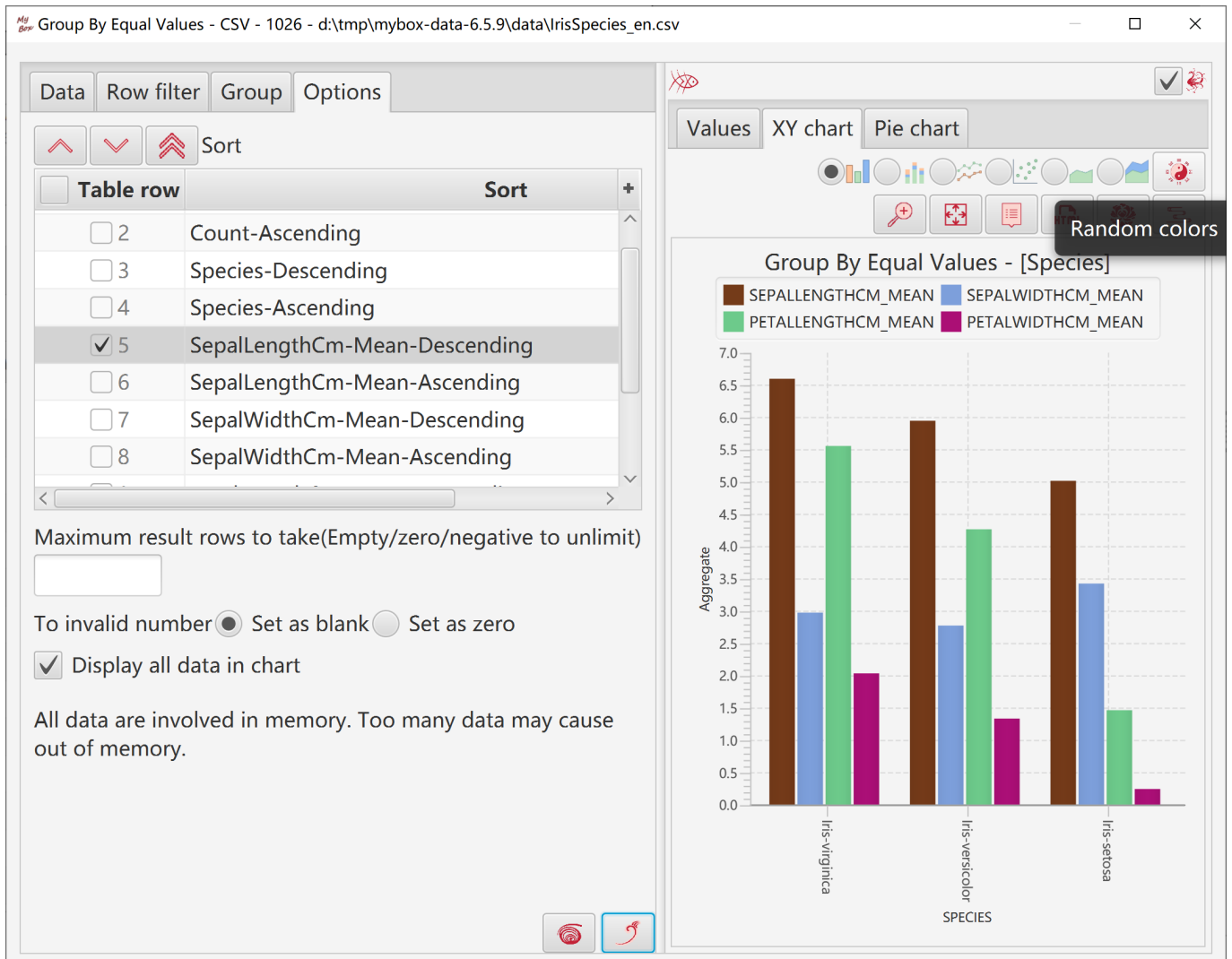
 

2.6.3 Group b Equal Values

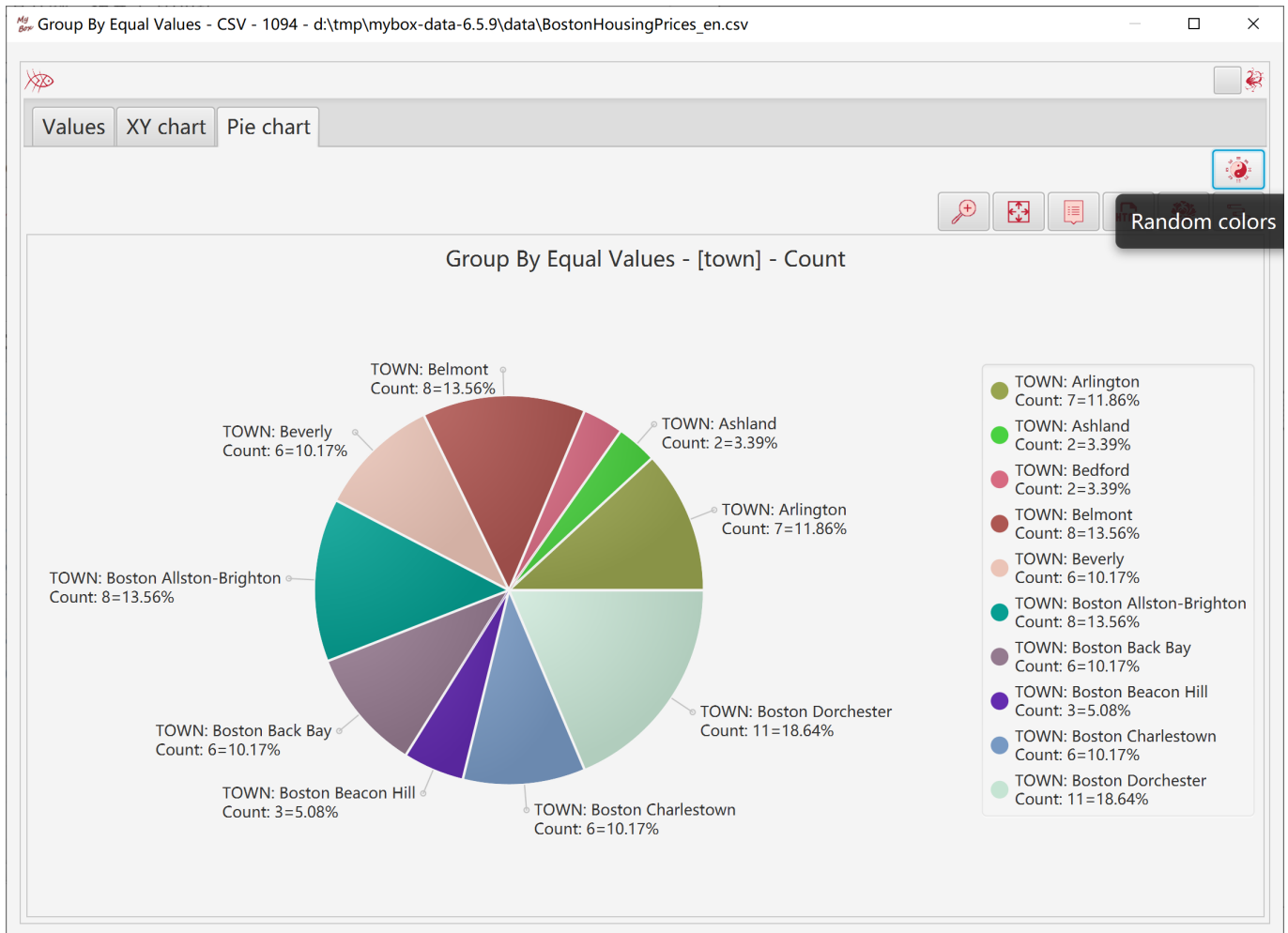
1. Select data:
 - Rows can be: current page, selected rows, or all pages.
 - Set row filter.
2. Calculate:
 - Select some columns as "group by". Need not be numbers.
 - Select some aggregate values to be calculated:
 - Like average, max, min, variance, standard deviation, etc.
 - "count" is always calculated, and not in the list.
 - Multiple items can be selected. None selection is permitted.
 - Involved values should be numbers.
 - Select some values as "order by". Multiple items can be selected. None selection is permitted
 - Input maximum number of result rows. Blank/zero/negative means no limitation.
 - To invalid numbers, option to set as blank or zero.

[illegible]

3. Calculated groups and their aggregate values are displayed in XY chart.



4. Calculated groups and their count are displayed in Pie chart.



5. In charts:
 - If only one column in group, category values are values of this column.
 - If multiple columns in group, category values are names of groups.
6. Option to display all data in charts:
 - If yes, lots of data may cause out of memory.
 - If no, only current page of results are display in charts.
7. Data overflow may happen.

2.6.4 Simple Linear Regression

2.6.4.1 Regression

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.

1. Select 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.
2. Calculate:
 - Select one column as independent variable.
 - Select another column as dependent variable.
 - Option whether includes intercept.
 - Set decimal scale.
3. 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.

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

Data Filter Options

Independent variable: income

Dependent variable: happiness

Desired significance level(alpha): 0.05

☒ Intercept(b0)

☒ Display all data in chart

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

About linear regression ☒ Decimal scale 4

Model Regression Fitting Residual

N...	income	happiness	Slope(b1)	Intercept(b0)	Coeffici...	Pearsc +
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

Selected: 0 Rows: 20/477 Page size 20 Page 1 /24

2.6.4.2 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 =

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

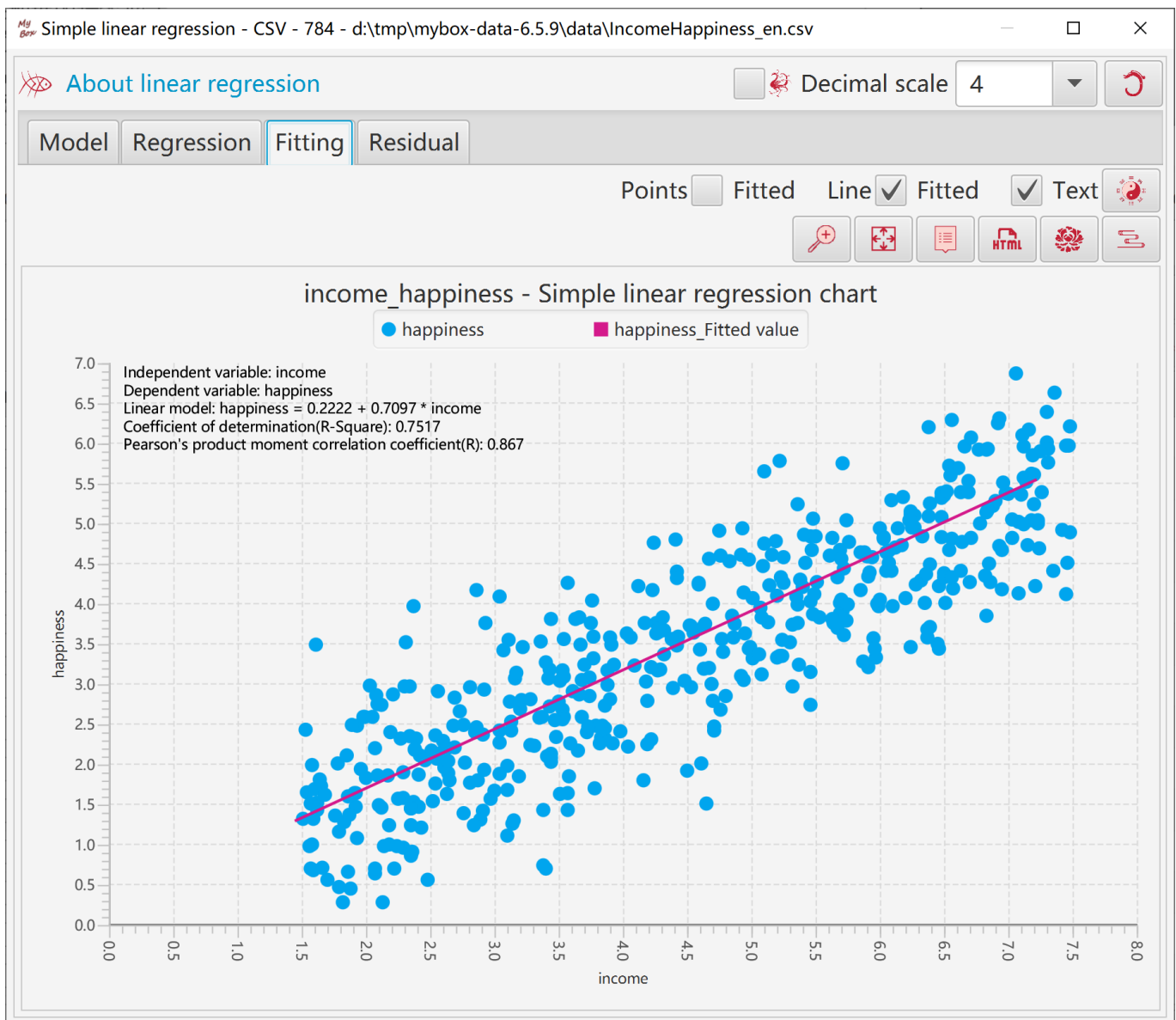
2.6.4.3 Fitted Chart

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.

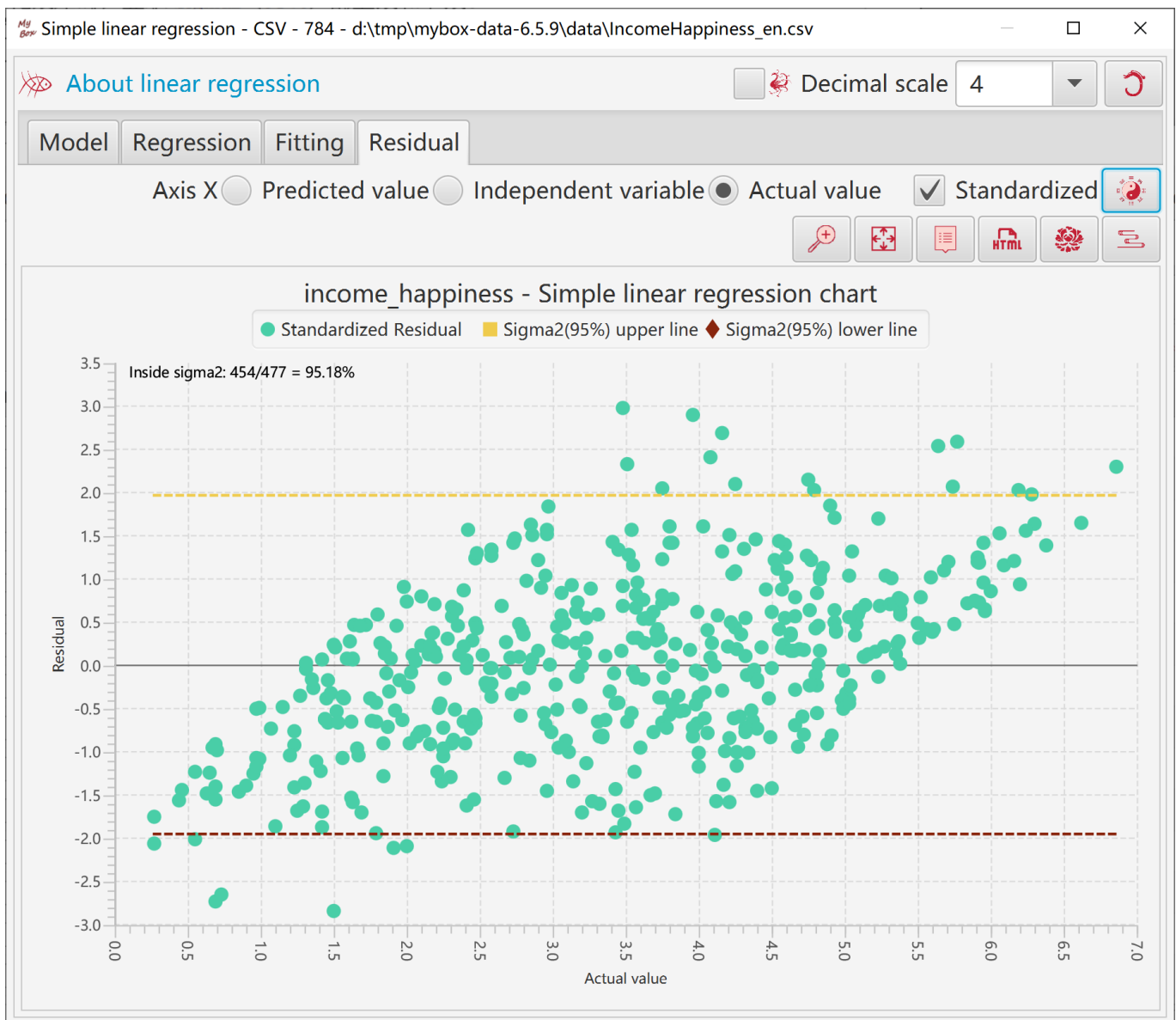


2.6.4.4 Residual Chart

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.



2.6.5 Simple Linear Regression - Combination

This tool helps to generate data of simple linear regression:

1. Select some columns.
2. Select options like decimal scale, alpha, whether include intercept.
3. When click button "OK":
 - Tool make pairs of columns and calculate their regression models.
 - Sort the modes by their coefficient of determination(R-Square) in descending order.
4. Select one mode and click button "View" to view its regression data, fitting chart, and residual chart.

MyBox Simple linear regression - Combination - CSV - 1026 - d:\tmp\mybox-data-6.5.9\data\IrisSpecies_en.csv

Data Filter Options

No selection means all columns

☒ SepalLengthCm
☒ SepalWidthCm
☒ PetalLengthCm
☒ PetalWidthCm ☐ Species

Decimal scale

Desired significance level(alpha)

☒ Intercept(b0)

About linear regression

Data row	Dependent va...	Independent ...	Coefficient of de...	Pearson's pr...	+
9	PetalLengthCm	PetalWidthCm	0.92690123	0.9627571	Peta
12	PetalWidthCm	PetalLengthCm	0.92690123	0.9627571	Peta
2	SepalLengthCm	PetalLengthCm	0.75995531	0.87175416	Sepa
7	PetalLengthCm	SepalLengthCm	0.75995531	0.87175416	Peta
3	SepalLengthCm	PetalWidthCm	0.66904815	0.81795363	Sepa
10	PetalWidthCm	SepalLengthCm	0.66904815	0.81795363	Peta
5	SepalWidthCm	PetalLengthCm	0.17683379	-0.4205161	Sepa
8	PetalLengthCm	SepalWidthCm	0.17683379	-0.4205161	Peta
6	SepalWidthCm	PetalWidthCm	0.12712369	-0.35654409	Sepa
11	PetalWidthCm	SepalWidthCm	0.12712369	-0.35654409	Peta
1	SepalLengthCm	SepalWidthCm	0.01196163	-0.10936925	Sepa
4	SepalWidthCm	SepalLengthCm	0.01196163	-0.10936925	Sepa

Double click selected item to view

2.6.6 Multiple Linear Regression

2.6.6.1 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.

Multiple linear regression - CSV - 1094 - d:\tmp\mybox-data-6.5.9\data\BostonHousingPrices_en.csv

Data Filter Options

Independent variable  

No selection means all columns

☐ town ☒ longitude ☒ latitude
☒ crimeratio ☒ zonedbigger_25000
☒ industrialland
☐ nearCharies_River
☒ nitrogendensity
☒ averageroom_number
☒ builtbefore_1940_ratio
☒ distanceto_centre
☒ accessibilityto_hightway ☒ taxrate
☒ pupilteacher_ratio
☒ lowerclass_ratio ☒ medianprice

Dependent variable

median_price ▼

Normalize ☒ L2(Z-Score)

Regression

☒ Ordinary Least Squares(OLS)
☒ Intercept(b0)

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

About linear regression  

Model Regression

Decimal scale 7   

Linear model: median_price = 0.0 - 0.0442545 * longitude + 0.0313181 * latitude - 0.1148961 * crime_ratio + 0.1193431 * zoned_bigger_25000 + 0.0174699 * industrial_land - 0.2009357 * nitrogen_density + 0.2768505 * average_room_number + 0.0265283 * built_before_1940_ratio - 0.3208416 * distance_to_centre + 0.293084 * accessibility_to_hightway - 0.2696422 * tax_rate - 0.2077848 * pupil_teacher_ratio - 0.4481128 * lower_class_ratio

Independent variable: longitude =

Independent variable: latitude =

Independent variable: crime_ratio =

Independent variable: zoned_bigger_25000 =

Independent variable: industrial_land =

Independent variable: nitrogen_density =

Independent variable: average_room_number =

Independent variable: built_before_1940_ratio =

Independent variable: distance_to_centre =

Independent variable: accessibility_to_hightway =

2.6.6.2 Model

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

Multiple linear regression - CSV - 1094 - d:\tmp\mybox-data-6.5.9\data\BostonHousingPrices_en.csv

About linear regression

Model Regression

Decimal scale 7

Independent variable: pupil_teacher_ratio =

Independent variable: lower_class_ratio = **Predict**

Dependent variable: median_price =

Model

Name	Value
Dependent variable	median_price
Independent variable	[longitude, latitude, crime_ratio, zoned_bigger_25000, industrial_land, nitrogen_density, average_room_number, built_before_1940_ratio, distance_to_centre, accessibility_to_hightway, tax_rate, pupil_teacher_ratio, lower_class_ratio]
Number of observations	506
Intercept(b0)	0.0
Coefficients	[-0.0442545, 0.0313181, -0.1148961, 0.1193431, 0.0174699, -0.2009357, 0.2768505, 0.0265283, -0.3208416, 0.293084, -0.2696422, -0.2077848, -0.4481128]
Coefficient of determination(R-Square)	0.7343658
Adjusted R squared	0.727347

TD

2.6.7 Frequency Distributions

1. Select 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.
2. Calculate:
 - Select one column to count frequency.
 - Option whether case-insensitive
 - Set decimal scale.
3. Target can be following: new csv/excel/text file, matrix, system clipboard, MyBox clipboard, database 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

Data Row filter Options

Column town ☒ Case-insensitive

Decimal scale 2

Target

External ☒ CSV ☐ XLS ☐ TXT ☐ Matrix ☐ Clipboard ☐ Database

Name BostonHousingPrices_en.csv

In table ☐ Insert above ☐ Append below ☐ Replace

2.6.8 Values Percentage

1. Select 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.
2. Calculate:
 - According to: Columns/rows/all.
 - Select how to treat negative values: zero or absolute value.
 - Set decimal scale.
 - To non-numeric, skip or count as zero.
3. Target can be following: new csv/excel/text file, matrix, system clipboard, MyBox clipboard, database table.
4. When rows are current page or selected ones, target can be defined location in table to insert/append/replace.

item	year 2020	year 2020_Perce...	year 2019	year 2019_Perce...
0	390.1	100	370.4	100
food consumption per capita(kilogram)	141.2	36.2	130.1	35.12
cooking oil consumption per capita(kilogram)	10.4	2.67	9.5	2.56
vegetables and mushrooms consumption per ca...	103.7	26.58	98.6	26.62
meat consumption per capita(kilogram)				
poultry consumption per capita(kilogram)				
aquatic products consumption per capita(kilog...				
eggs consumption per capita(kilogram)				
milk consumption per capita(kilogram)				
fruits consumption per capita(kilogram)				
sugar consumption per capita(kilogram)				

Value percentage - CSV - 1102 - d:\tmp\mybox-data-6.5.9\data\ChinaF...

Percentage

Data

Row filter

Options

Columns   No selection means all columns

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

According to ☒ Columns ☐ Rows ☐ All

☒ Data values

Decimal scale

To negative ☒ Skip ☐ Count as zero ☐ Absolute value

To invalid number ☒ Skip ☐ Count as zero

Target

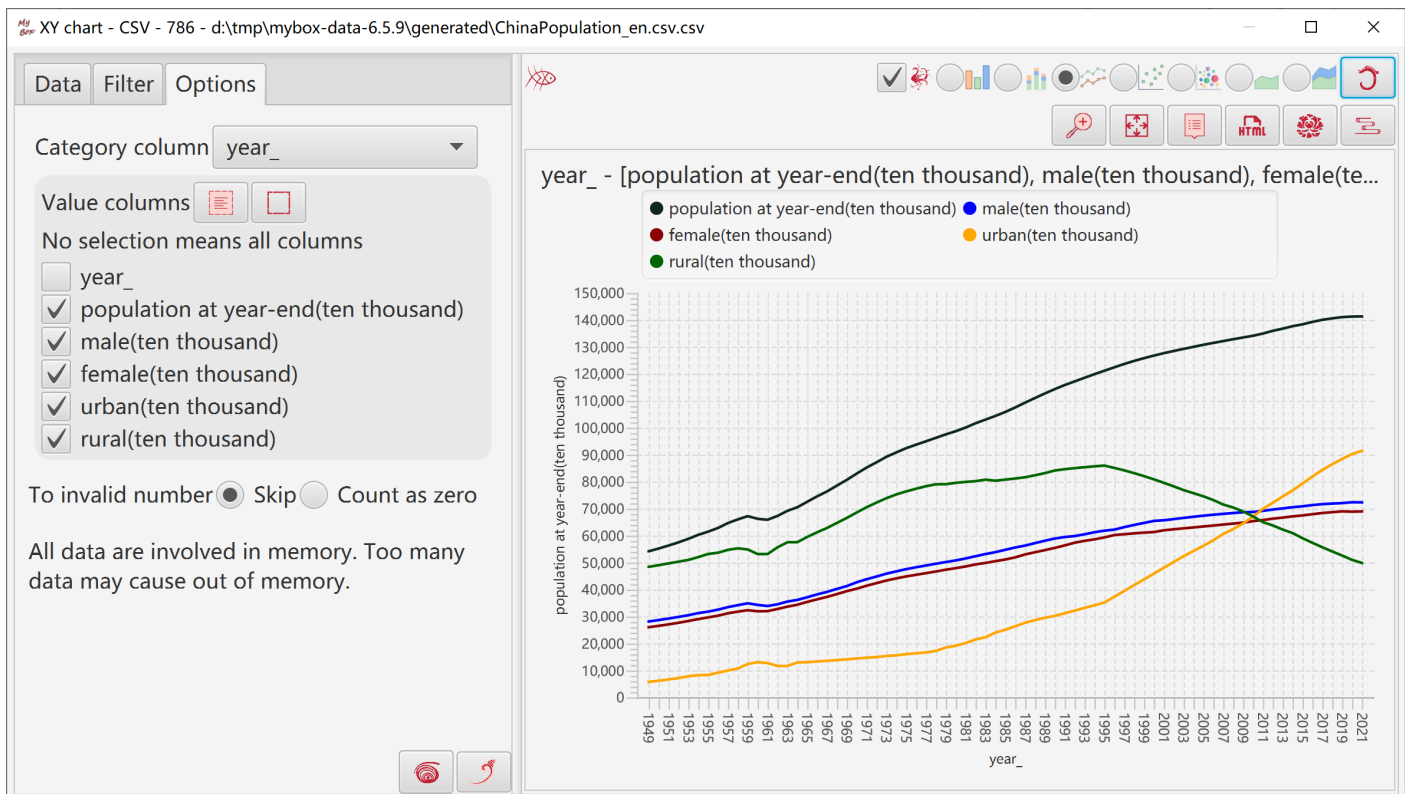
External ☒ CSV ☐ XLS ☐ TXT ☐  ☐  ☐  ☐  ☐ 

2.7 Chart

2.7.1 XY Chart

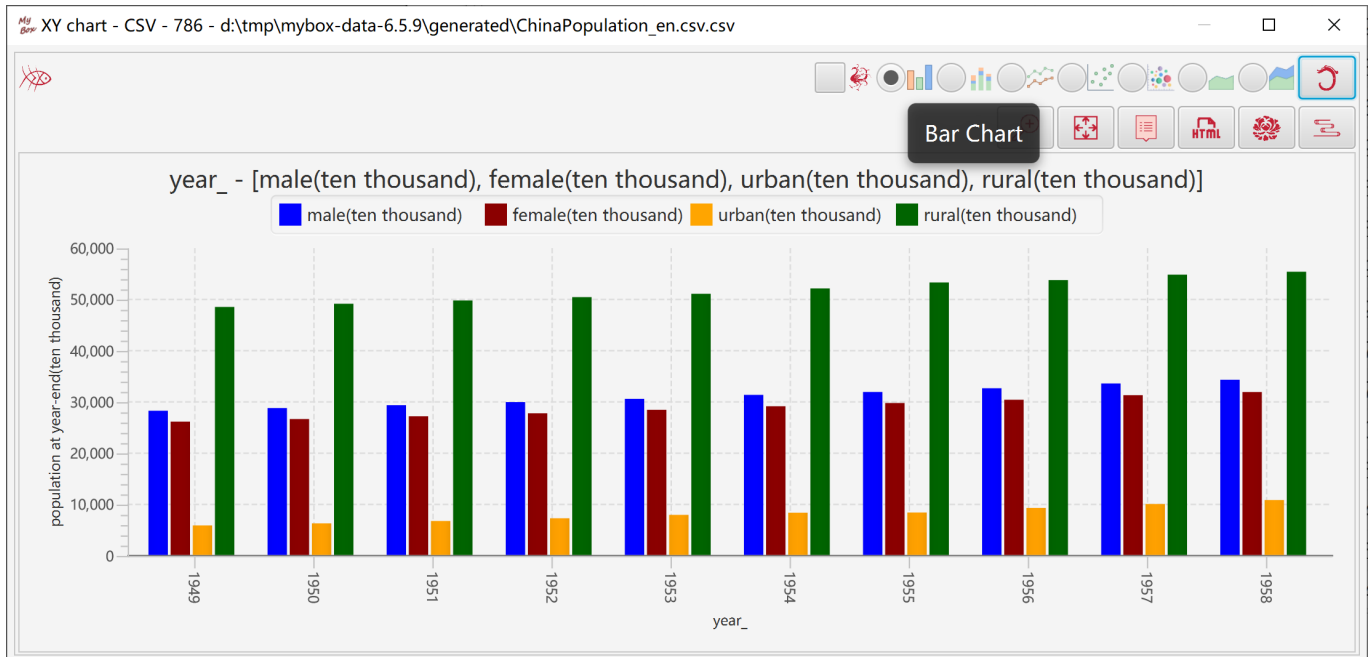
2.7.1.1 Data

1. Select 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.
2. Calculate:
 - Select one column as "Category Axis", to define data names..
 - Multiple columns can be selected as data in direction of "Number Axis". Different value series are shown in different colors or shapes.
 - To non-numeric, skip or count as zero.
 - When display all data rows(all pages), need concern memory limitation.
 - By default, "Category Axis" is the horizontal axis and "Number Axis" is the vertical axis.
3. Select type of XY Chart.
4. Click button “Menu” to set parameters of chart.
5. Click button “Pop” to display current chart as image in popped window.
6. Click button “Data” to display data of the chart in data table.
7. Click button “Html” to display data of the chart in html page.



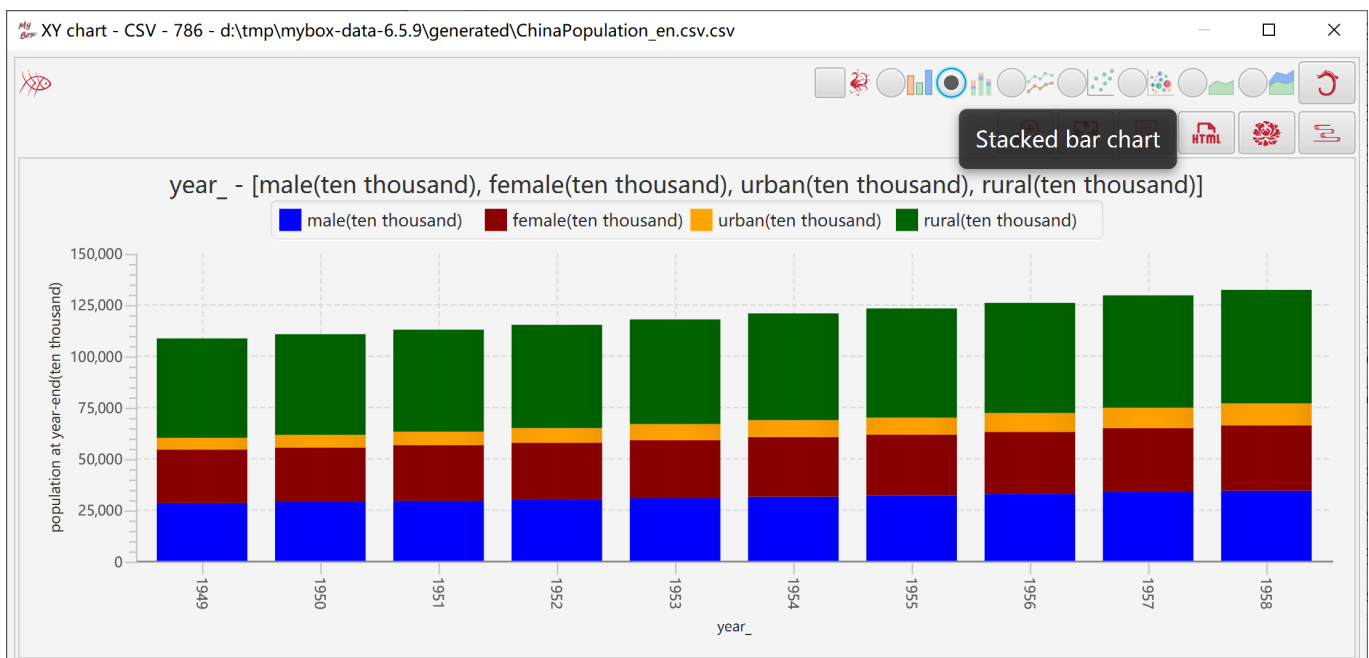
2.7.1.2 Bar Chart

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



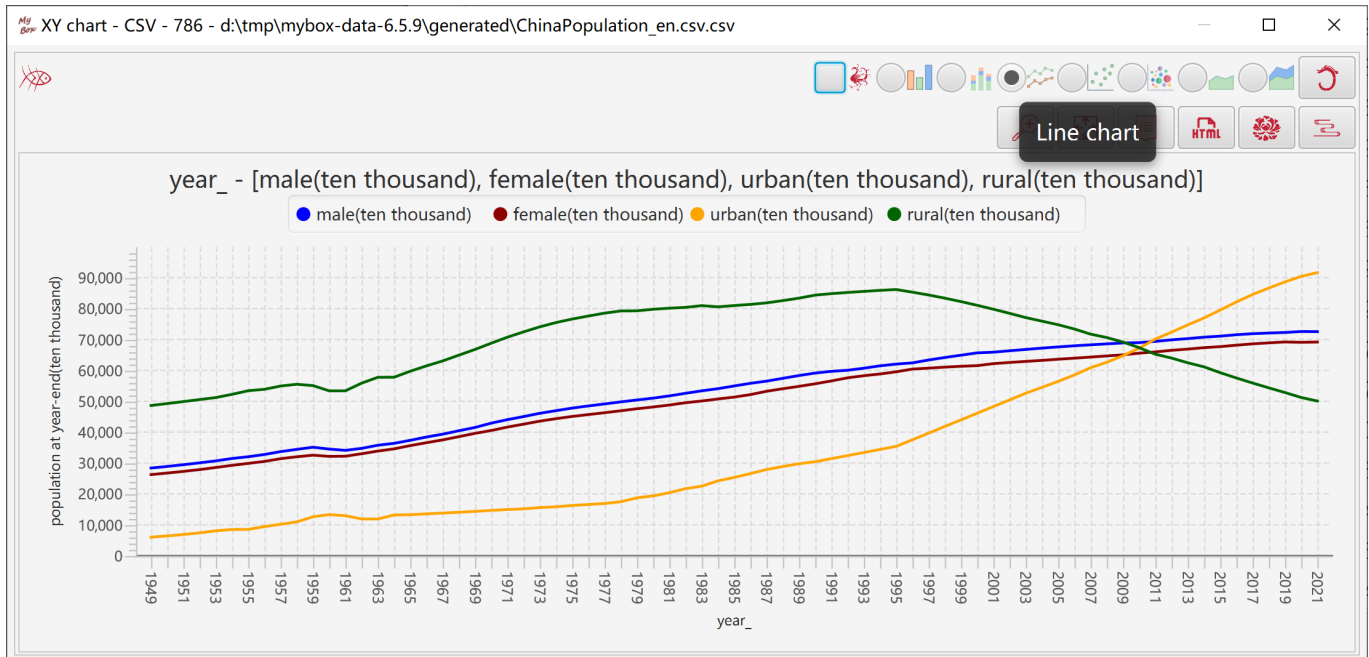
2.7.1.3 Stacked Bar Chart

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



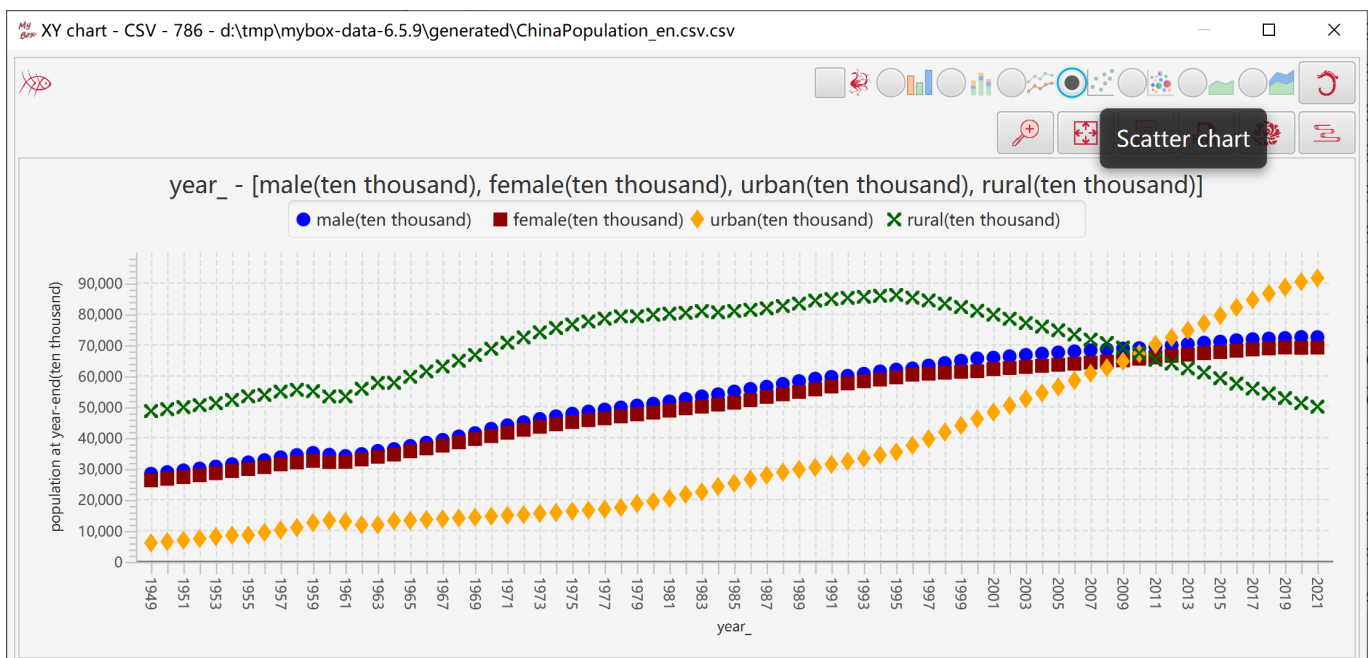
2.7.1.4 Line Chart

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



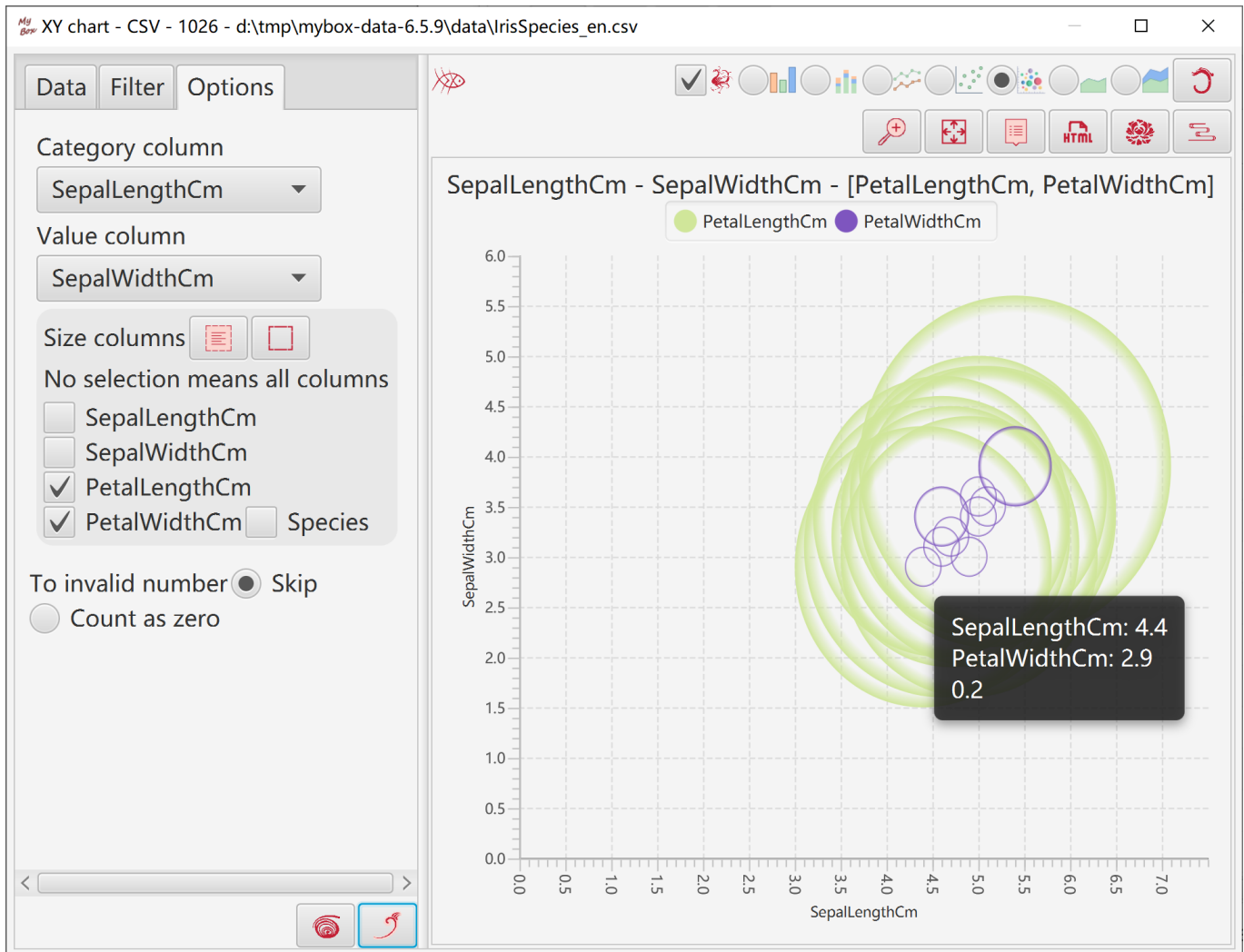
2.7.1.5 Scatter Chart

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



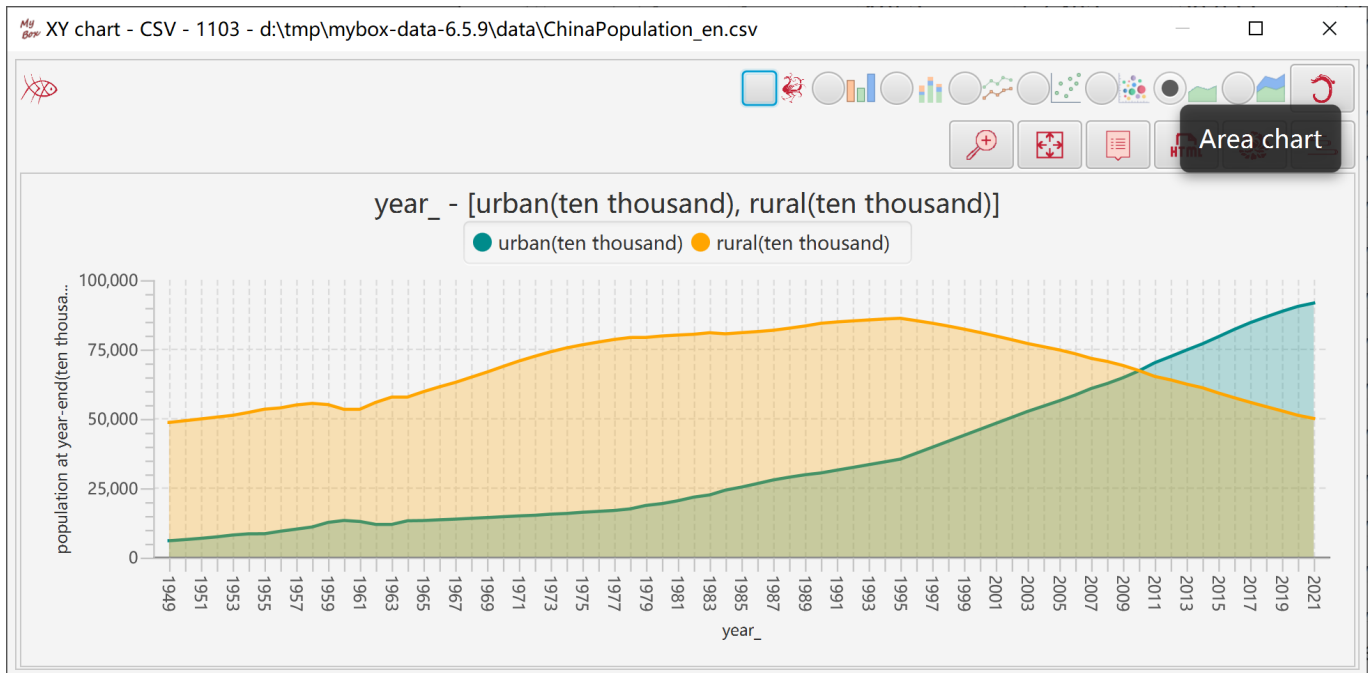
2.7.1.6 Bubble Chart

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.



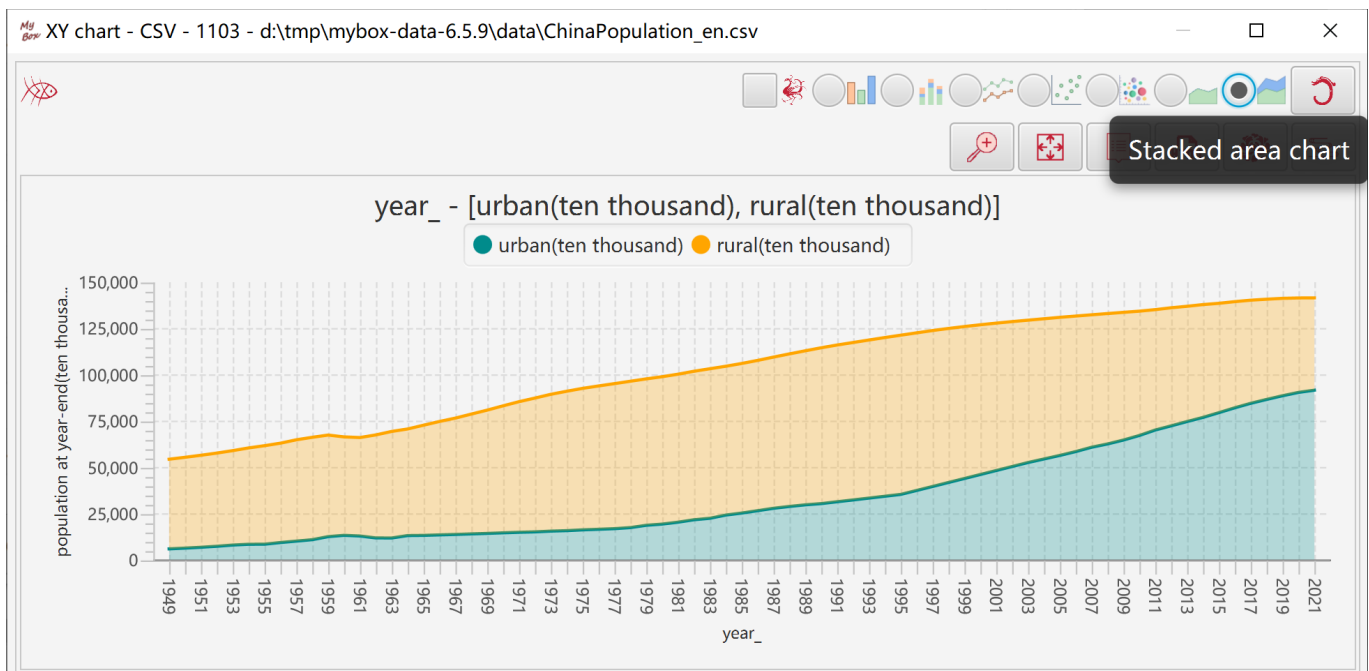
2.7.1.7 Area Chart

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



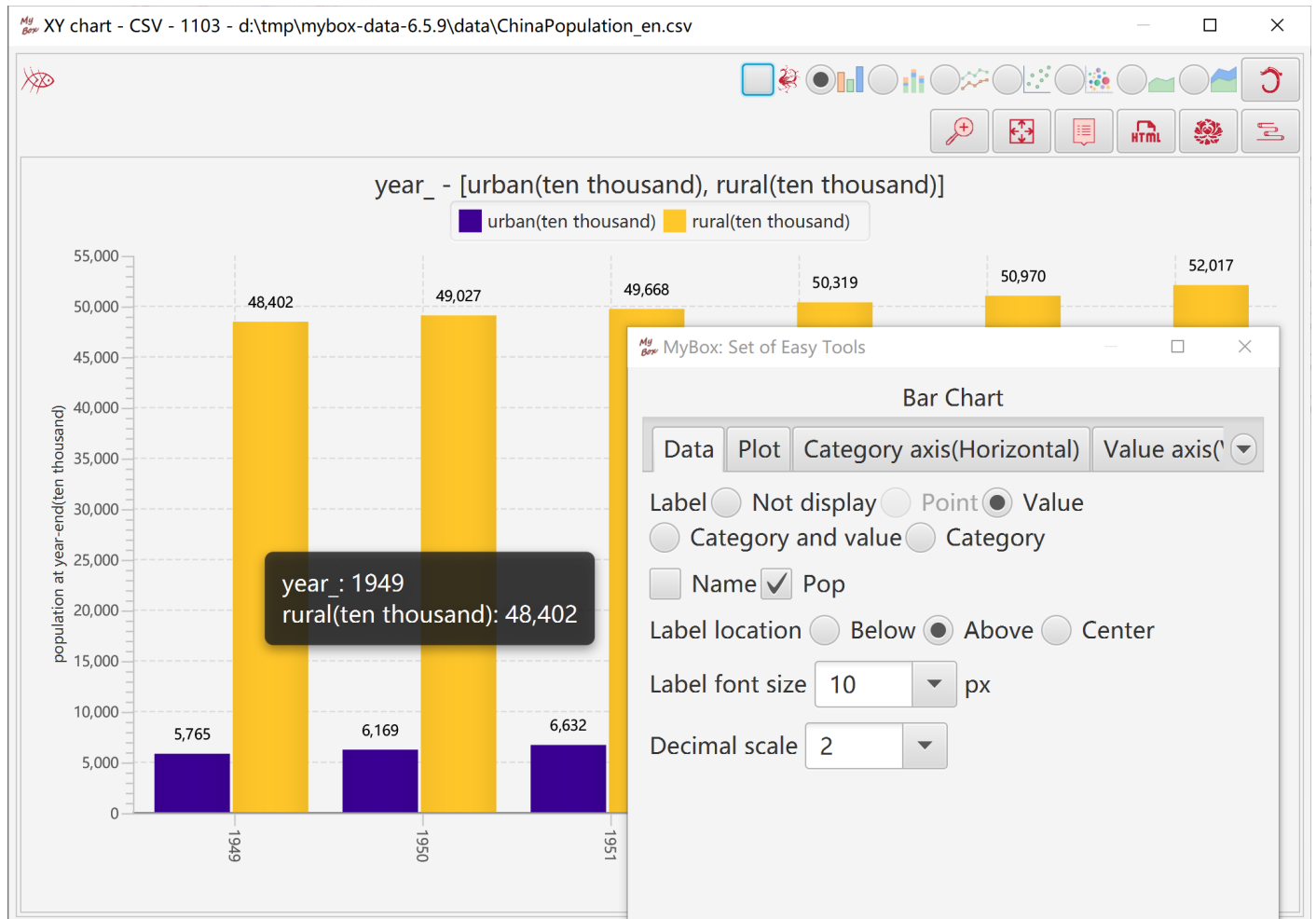
2.7.1.8 Stacked Area Chart

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



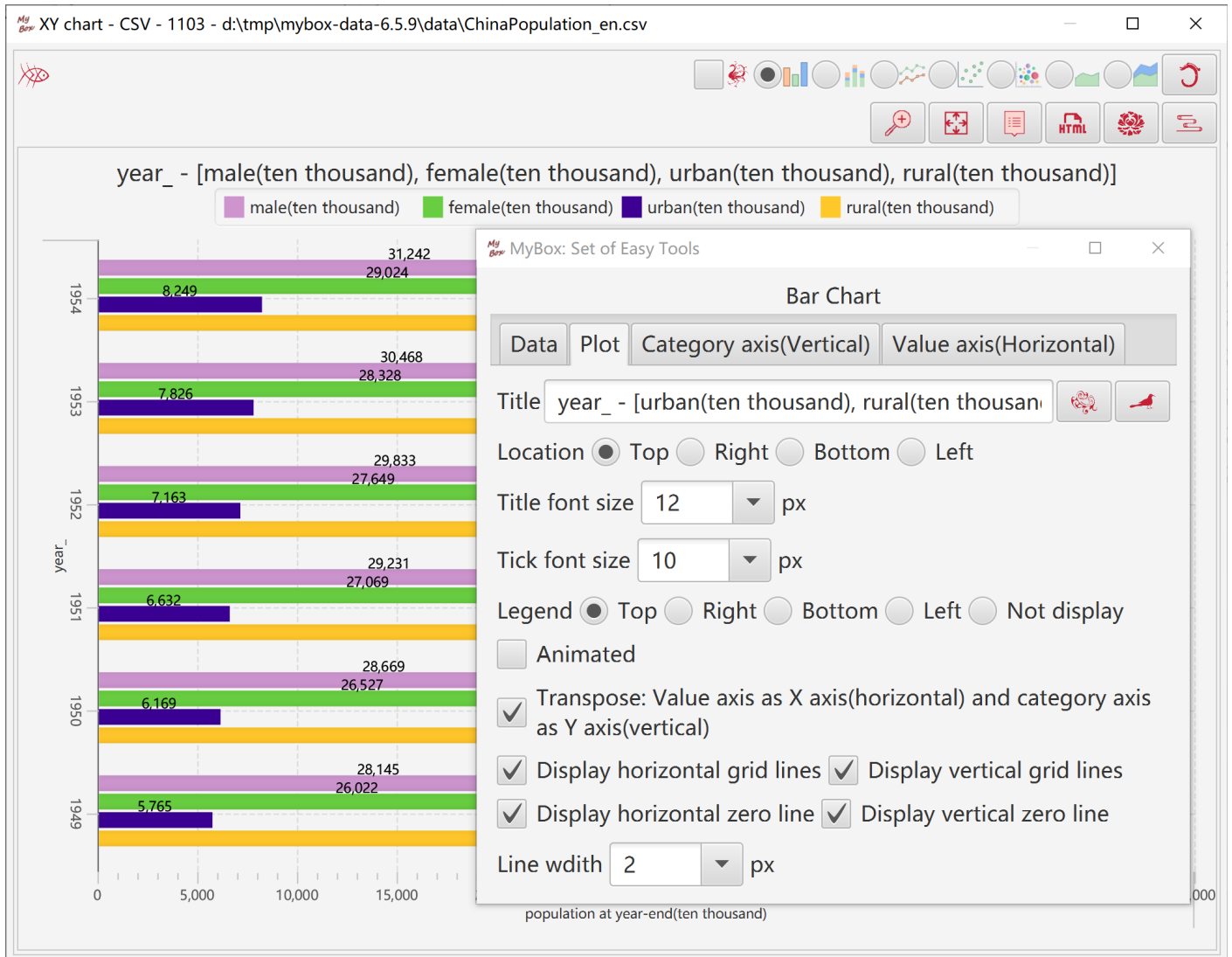
2.7.1.9 Parameters of Data in Chart

1. Options about label: not display, point, value, category, etc.
2. Location of labels.
3. Font size of labels.
4. Decimal scale.



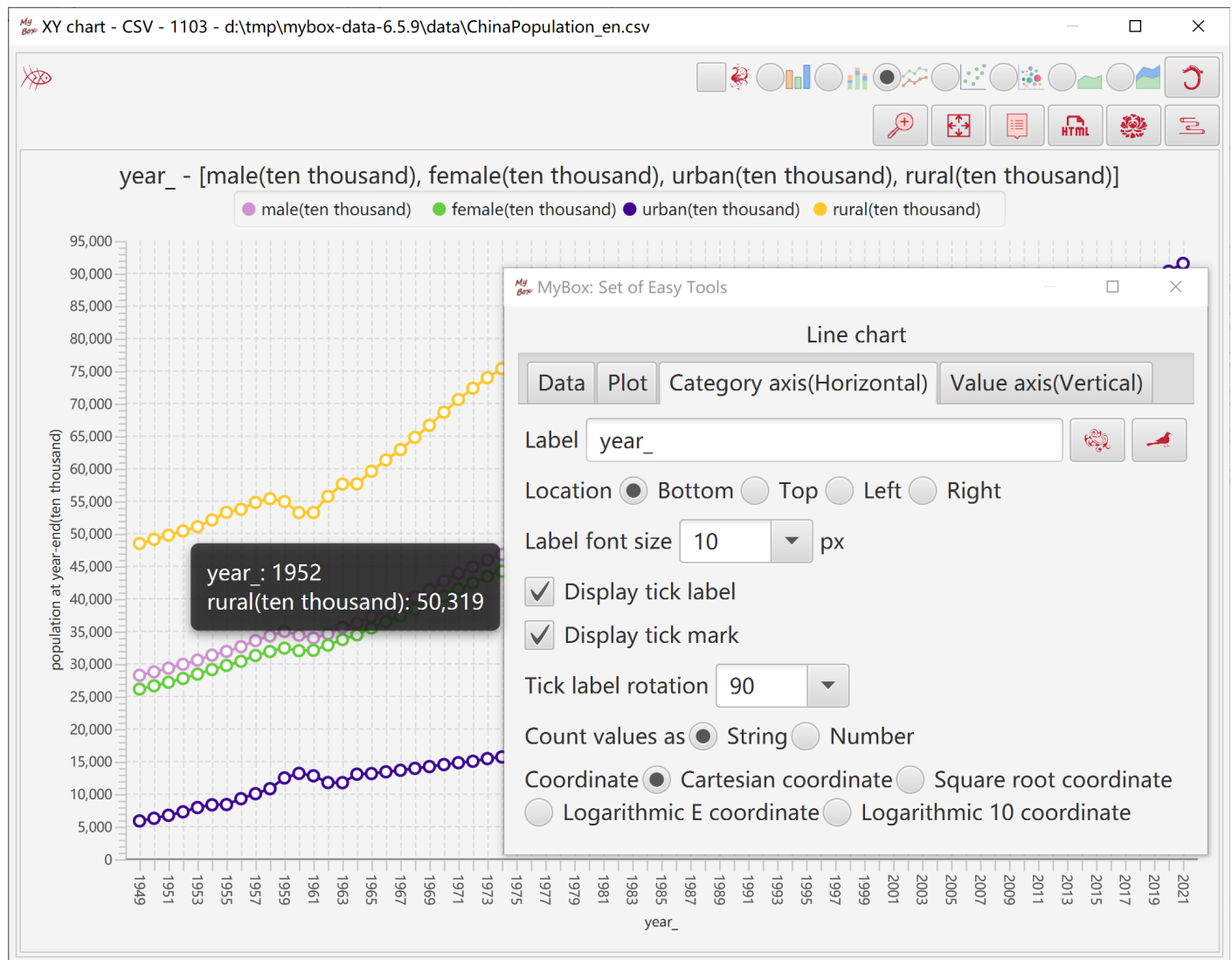
2.7.1.10 Layout

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



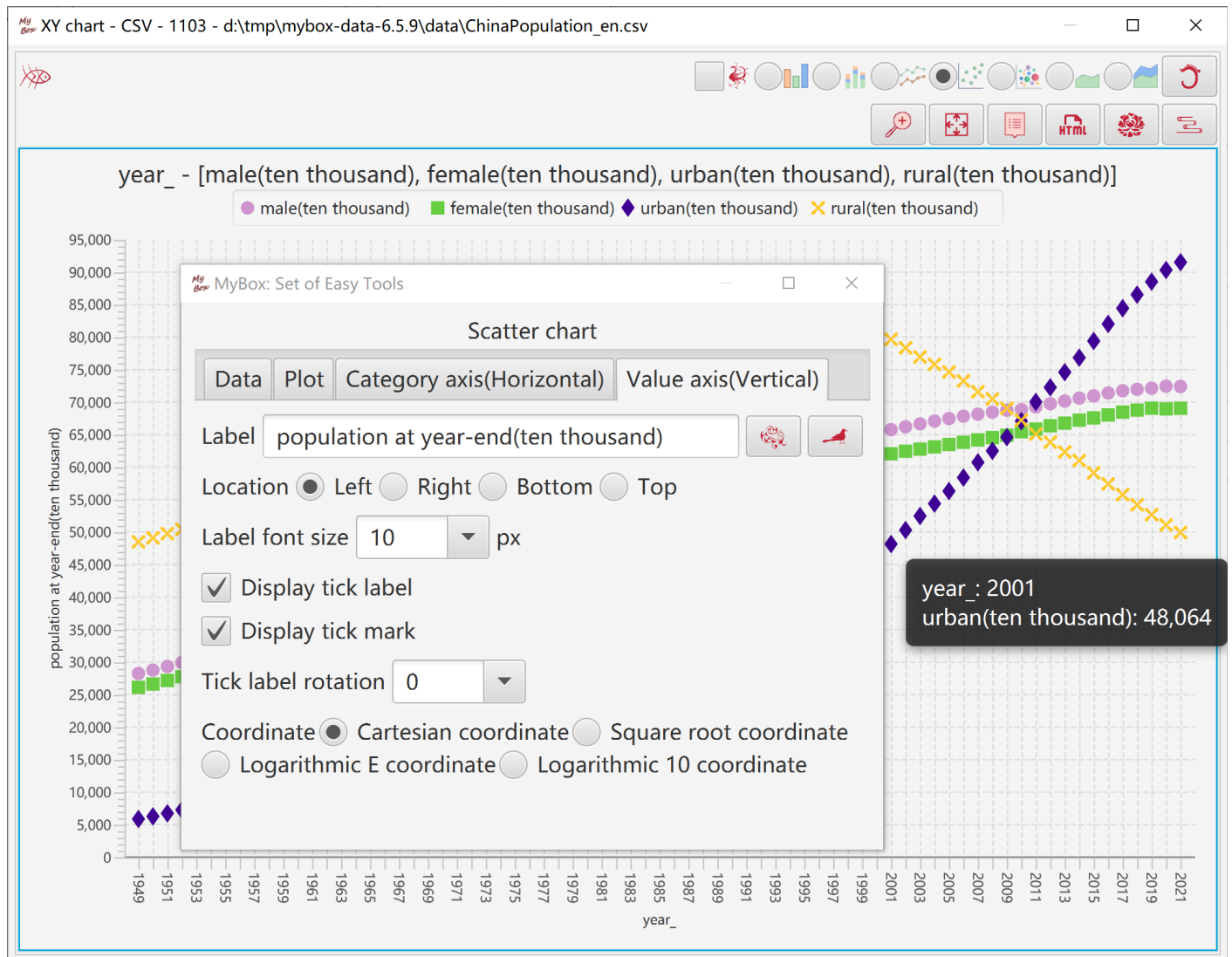
2.7.1.11 Category Axis

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



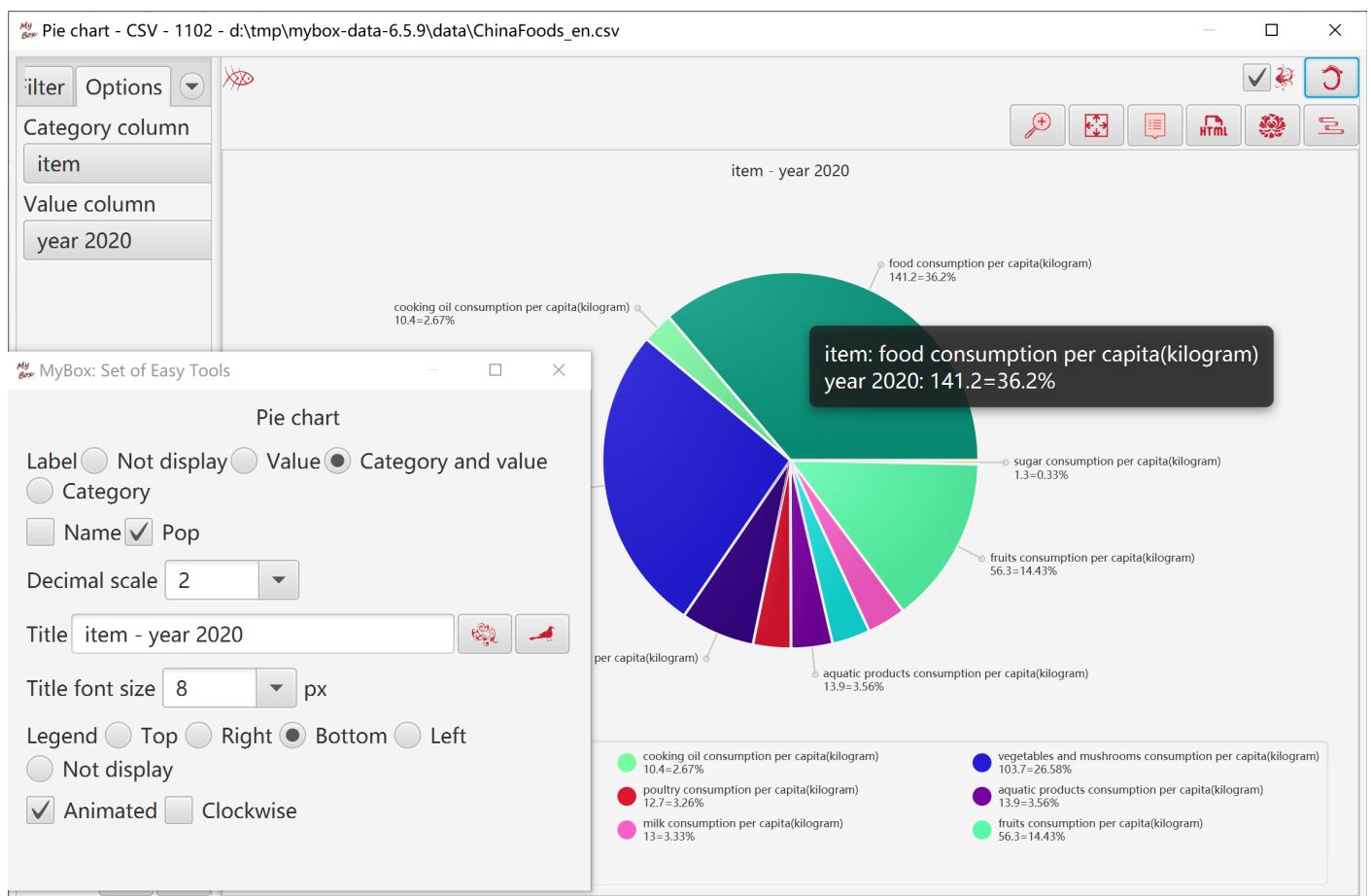
2.7.1.12 Number Axis

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



2.7.2 Pie Chart

1. Select 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.
2. Calculate:
 - Select one column as "Category Axis", to define data names..
 - Select another column as "Number Axis".
 - Data numbers are represented as percentages with a circle divided into segments.
 - Value column should be non-negative.
 - When display all data rows(all pages), need concern memory limitation.
3. Click button “Menu” to set parameters of chart.
4. Click button “Pop” to display current chart as image in popped window.
5. Click button “Data” to display data of the chart in data table.
6. Click button “Html” to display data of the chart in html page.



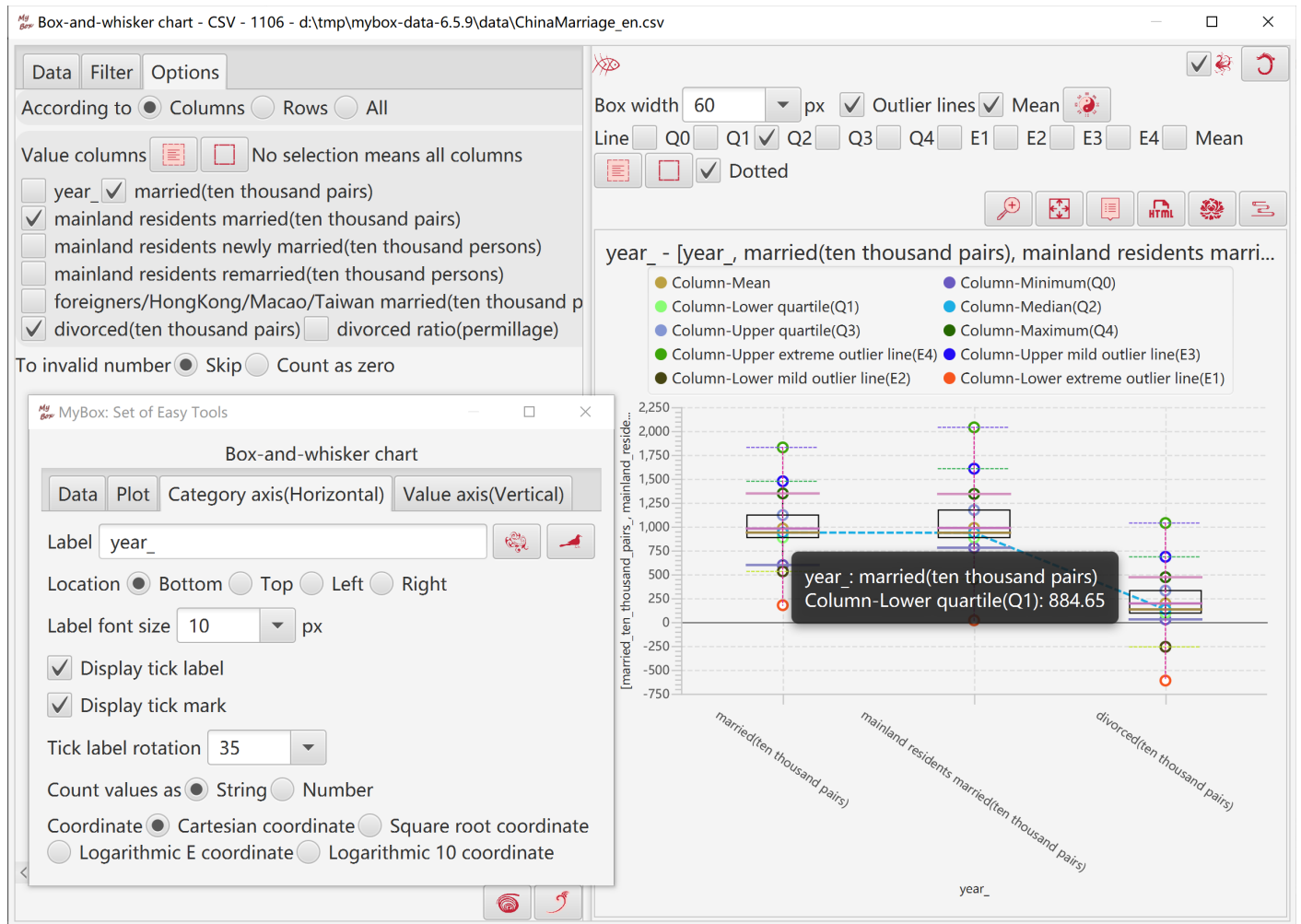
2.7.3 Box-and-whisker Chart

1. Box-and-whisker chart represents data distribution:
 - Sort data according to column/row/all in ascending order.
 - Following items can show aggregation and discreteness of data:

Minimum	$Q0 =$ in 0%(start) of the data list
Lower quartile	$Q1 =$ in 25% of the data list
Median	$Q2 =$ in 50%(middle) of the data list
Upper quartile	$Q3 =$ in 75% of the data list
Maximum	$Q4 =$ in 100%(end) of the data list
 - Following values can be used to mark outliers of data:

Lower extreme outlier line	$E1 = Q1 - 3 * (Q3 - Q1)$
Lower mild outlier line	$E2 = Q1 - 1.5 * (Q3 - Q1)$
Upper mild outlier line	$E3 = Q3 + 1.5 * (Q3 - Q1)$
Upper extreme outlier line	$E4 = Q3 + 3 * (Q3 - Q1)$
 - Following values can be referred for discreteness:

Mean = average of the data list
2. Calculate:
 - Based on line chart.
 - According to: columns, rows(select category column), or all.
3. Set or select:
 - Box width.
 - Whether display outliers lines or mean.
 - Whether display connection lines of values, and whether dotted lines.
 - Random colors.
4. Click button “Menu” to set parameters of chart.
5. Click button “Pop” to display current chart as image in popped window.
6. Click button “Data” to display data of the chart in data table.
7. Click button “Html” to display data of the chart in html page.



2.7.4 Self Comparison Bars Chart

1. Self comparison bars show difference between data and reference values.

Bars are calculated by following rulers:

- If value is zero, no bar
- When compare as absolute values:

$$\text{maximum_value} = \text{maximum_absolute_value_of_column/row/all}$$

$$\text{percentage} = \text{absolute_value} / \text{maximum_value}$$

$$\text{width} = \text{maximum_width} * \text{percentage}$$

$$\text{color} = \text{If value is larger than zero, color_of_column.}$$

$$\text{If value is less than zero, inverted_color_of_column}$$

- When compare as range of minimum and maximum:

$$\text{maximum_value} = \text{maximum_value_of_column/row/all}$$

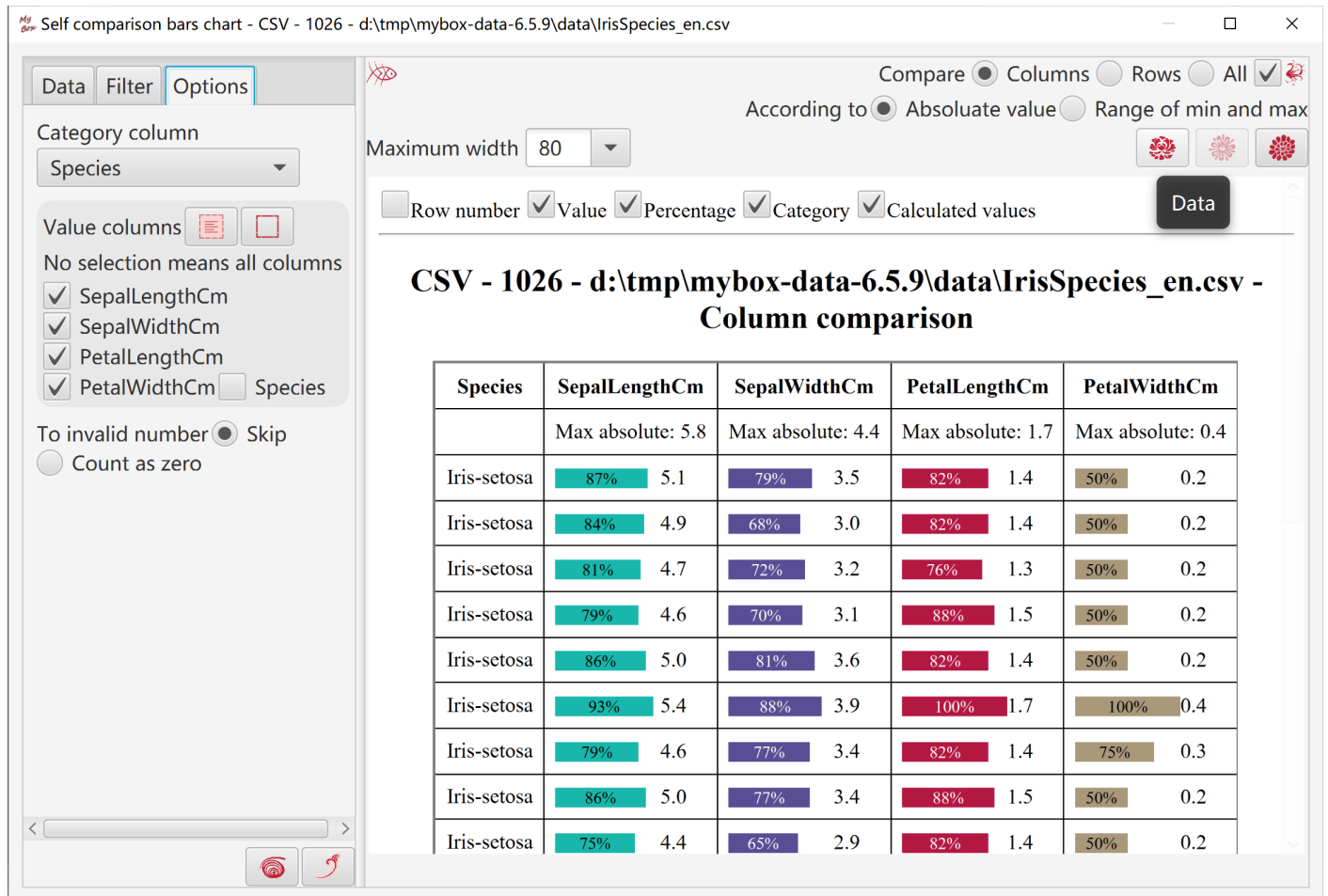
$$\text{minimum_value} = \text{minimum_value_of_column/row/all}$$

$$\text{percentage} = (\text{value} - \text{minimum_value}) / (\text{maximum_value} - \text{minimum_value})$$

$$\text{width} = \text{maximum_width} * \text{percentage}$$

$$\text{color} = \text{color_of_column}$$

2. Data: Select rows in table or select all data rows(all pages), and select columns.
3. Calculate:
 - Select one column as category column(unnecessary)
 - Compare: columns/rows/all.
 - According to: absolute values, or range of minimum and maximum.
 - When display all data rows(all pages), need concern memory limitation.
4. Set or select:
 - Maximum width
 - Whether display row numbers, values, percentages, categories, calculated values.
5. Edit data in chart.
6. Edit html of chart.

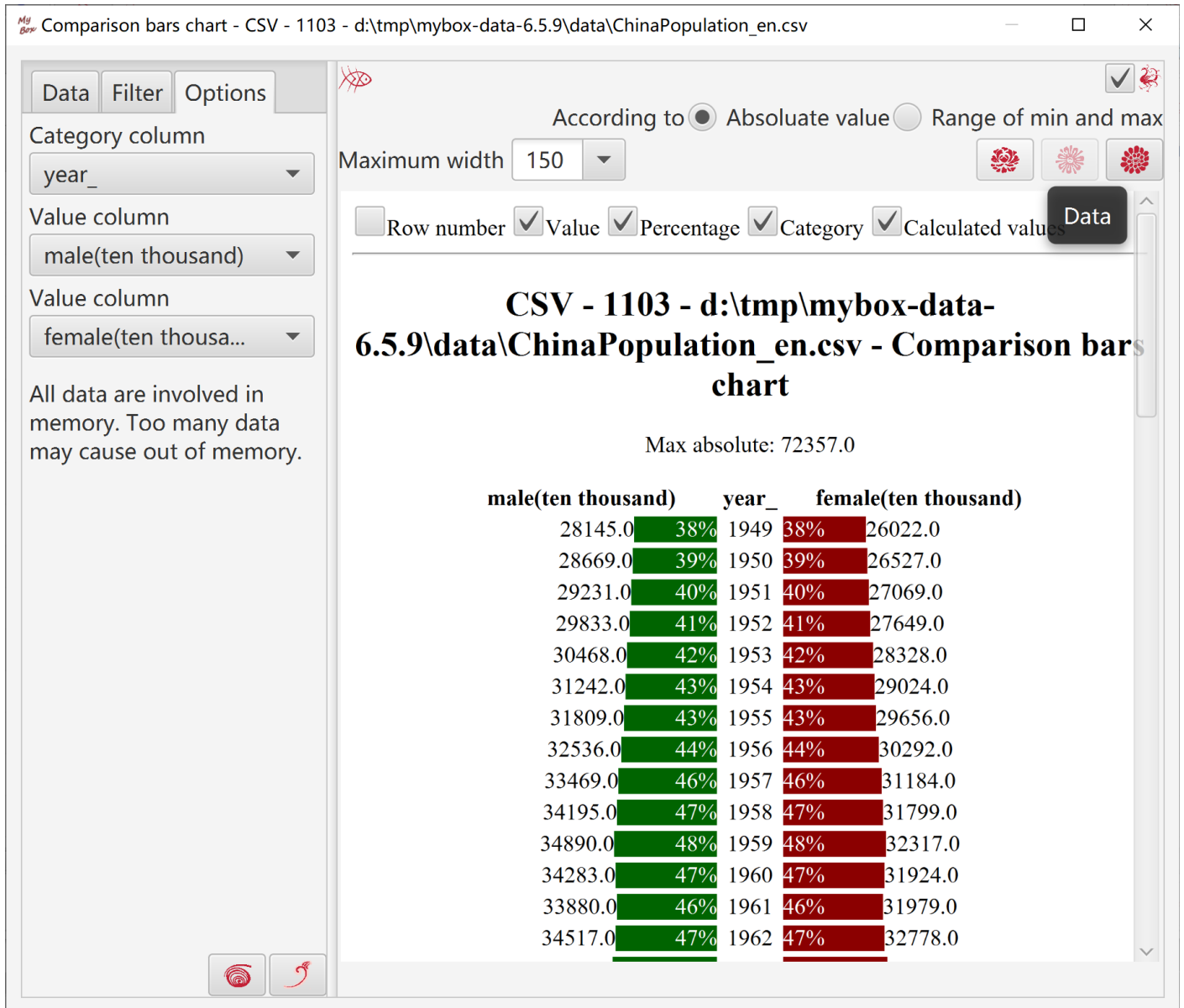


2.7.5 Comparison Bars Chart

1. Comparison bars show difference between two series of data.

Bars are calculated by following rulers:

- If value is zero, no bar
 - When compare as absolute values:
 - $\text{maximum_value} = \text{maximum_absolute_value_of_value_columns}$
 - $\text{percentage} = \text{absolute_value} / \text{maximum_value}$
 - $\text{width} = \text{maximum_width} * \text{percentage}$
 - $\text{color} = \text{If value is larger than zero, color_of_column.}$
 - $\text{If value is less than zero, inverted_color_of_column}$
 - When compare as range of minimum and maximum:
 - $\text{maximum_value} = \text{maximum_value_of_value_columns}$
 - $\text{minimum_value} = \text{minimum_value_of_value_columns}$
 - $\text{percentage} = (\text{value} - \text{minimum_value}) / (\text{maximum_value} - \text{minimum_value})$
 - $\text{width} = \text{maximum_width} * \text{percentage}$
 - $\text{color} = \text{color_of_column}$
2. Data: Select rows in table or select all data rows(all pages), and select columns.
 3. Calculate:
 - Select one column as category column(unnecessary)
 - Select tow value columns
 - According to: absolute values, or range of minimum and maximum.
 - When display all data rows(all pages), need concern memory limitation.
 4. Set or select:
 - Maximum width
 - Whether display row numbers, values, percentages, categories, calculated values.
 5. Edit data in chart.
 6. Edit html of chart.



2.8 Manage Data

This tool manages following objects:

1. Data Files

- Record is created/updated when csv/excel/texts data file is opened by its editor.
- Data are saved in data file.
- Deleting record of data file will not delete data file itself.

2. Data Clipboards

- Record is created when data is copied into MyBox Clipboad.
- Data are saved in file under MyBox internal path.
- Deleting record of data clipboards will delete its internal file.

3. Matrices

- Records are maintained by Matrices Manager.
- Data are saved in MyBox database.
- Deleting record of matrix will delete data of this matrix.

4. Data Tables

- Records are maintained by Data Tables Manager.
- Data are saved in MyBox database tables.
- Deleting record of data table will delete data of this data table.

The screenshot shows the 'Manage Data' window for a database table named '1081 - CHINAPOPULATION_EN'. The window has a menu bar (Window, Document, Image, Data, File, Media, Network, Settings, Recent Accessed, Development, Help) and a toolbar with icons for opening, saving, and deleting files. A list of data objects is displayed on the left, with columns for 'Row...', 'Colu...', and 'Data row'. The selected object is '1081 - CHINAPOPULATION_EN' (Database Table). A 'Query' dialog is open, showing options for 'Type' (CSV, XLS, TXT, etc.) and 'Order by' (Descending, Modify time, Name, ID, Rows number, Columns number, File). The dialog also shows 'Selected: 0', 'Rows: 50/73', and 'Page size: 50'.

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

2.9 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.

Splice Data : Database Table - 1081 - CHINAPOPULATION_EN

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

Tab...	ID	Type	Set as data B	Row...	Colu...
☐ 1	1105	Texts	ChinaPo...	73	6
☐ 2	1061	Texts	ChinaPo...	73	6
☐ 3	1089	MyBox Clipbo...	ChinaPo...	6	7
☐ 4	1086	MyBox Clipbo...	ChinaPo...	6	7
☐ 5	1085	MyBox Clipbo...	ChinaPo...	6	7
☐ 6	1083	MyBox Clipbo...	ChinaPo...	1	7
☐ 7	1082	MyBox Clipbo...	ChinaPo...	6	7
☐ 8	1063	MyBox Clipbo...	a	6	3
☐ 9	1065	Matrix	a	3	3
☐ 10	1064	Matrix	b	6	4
☐ 11	1062	Excel	ChinaPo...	73	6
☒ 12	1081	Database Table	ChinaPo...	73	7
☐ 13	1069	Database Table	ed	3	3
☐ 14	1068	Database Table	ChinaPo...	73	7
☐ 15	1103	CSV	ChinaPo...	73	6

Page 1 / 10 Page size 50 Rows: 50/457 Selected: 1

Splice Data

Direction
☒ Horizontal ☐ Vertical

Rows number
 By ☒ Data A ☐ Data B ☐ Longer ☐ Shorter

Target
☒ CSV ☐ XLS ☐ TXT ☐ [Matrix] ☐ [Clipboard] ☐ [Database]

Name

2.10 Data File

2.10.1 CSV File

1. When file is loaded abnormally, change options and click Refresh button.
2. Options: charset, whether has first line as field names, and delimiter of data.

MyBox Edit CSV File : CSV - 1103 - d:\tmp\mybox-data-6.5.9\data\ChinaPopulation_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

CSV delimiter

☒ , ☐ | ☐ @ ☐ # ☐ ; ☐ Character

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

Backup

Save as

Charset UTF-8

☒ First line defines columns names

CSV delimiter ☒ , ☐ | ☐ @ ☐ # ☐ ;

☐ Character

☐ Load after save as

☒ Open after save as

☐ Just save as

Columns Attributes View Edit

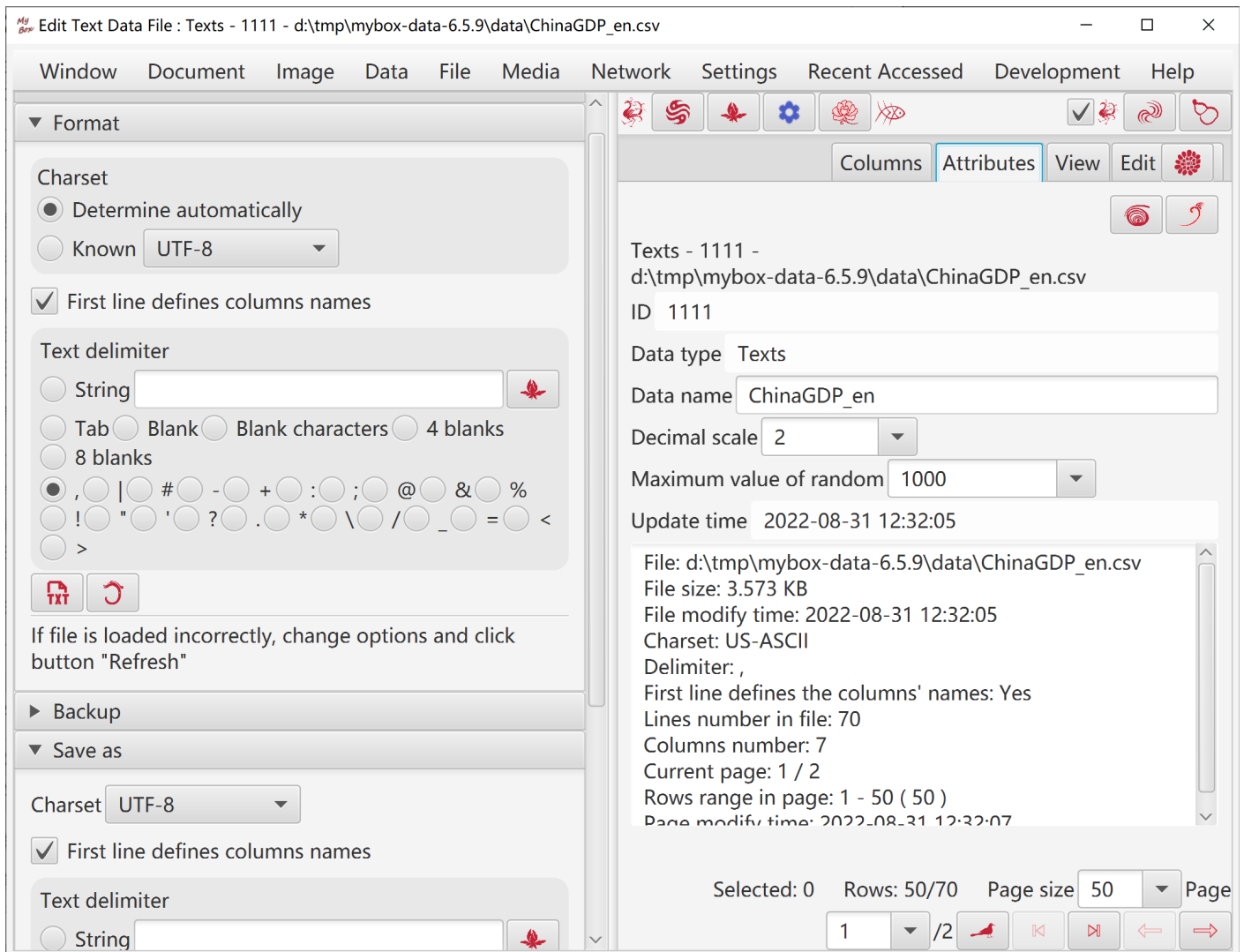
Table row Data row year_ populatio... male(ten ... femal +

☐ 1	1	1949	54,167	28,145	26,022
☐ 2	2	1950	55,196	28,669	26,527
☐ 3	3	1951	56,300	29,231	27,069
☐ 4	4	1952	57,482	29,833	27,649
☐ 5	5	1953	58,796	30,468	28,328
☐ 6	6	1954	60,266	31,242	29,024
☐ 7	7	1955	61,465	31,809	29,656
☐ 8	8	1956	62,828	32,536	30,292
☐ 9	9	1957	64,653	33,469	31,184
☐ 10	10	1958	65,994	34,195	31,799
☐ 11	11	1959	67,207	34,890	32,317
☐ 12	12	1960	66,207	34,283	31,924
☐ 13	13	1961	65,859	33,880	31,979
☐ 14	14	1962	67,296	34,517	32,778
☐ 15	15	1963	69,172	35,533	33,639

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

2.10.2 Text File

1. When file is loaded abnormally, change options and click Refresh button.
2. Options: charset, whether has first line as field names, and delimiter of data.



2.10.3 Excel File

- Options include sheet number and whether has first line as field names.
- Add/Delete/Rename sheets.
- 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.

MyBox Edit Excel File : Excel - 1112 - d:\tmp\mybox-data-6.5.9\generated\ChinaGraduates_en.xlsx - Sheet1

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

Information

Format

☒ First line defines columns names



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

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

Sheet

Sheet1 

Backup

Save as

☒ First line defines columns names

☐ Save current sheet only

☐ Load after save as

☒ Open after save as

☐ Just save as



Columns Attributes View Edit

 ##     

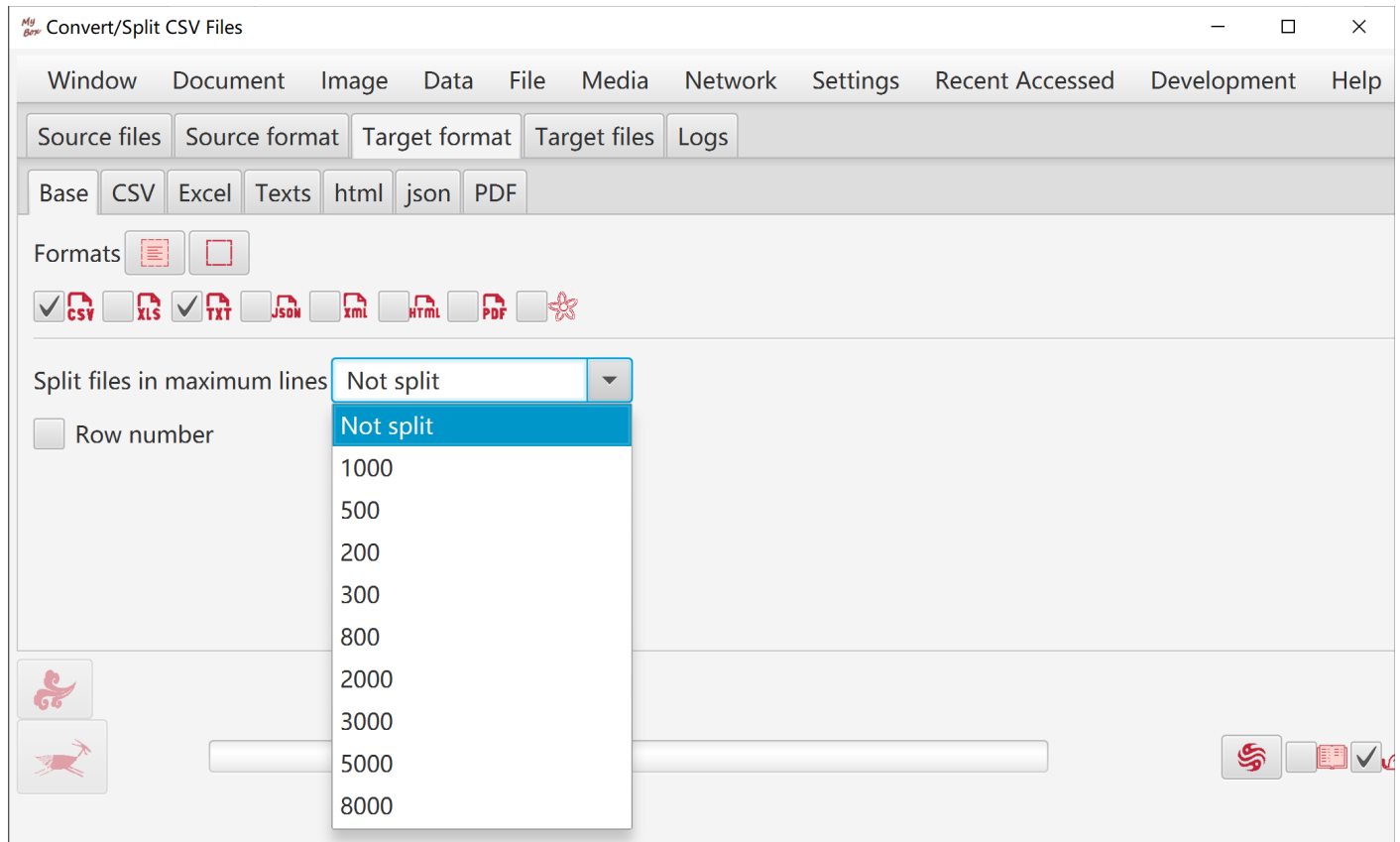
Tab...	Data row	year_	college gr...	middle sc...	high s...
☐ 1	1	1949	2.1	28	6.1
☐ 2	2	1950	1.8	29.6	6.2
☐ 3	3	1951	1.9	28.4	5.9
☐ 4	4	1952	3.2	22.1	3.6
☐ 5	5	1953	4.8	45.4	5.6
☐ 6	6	1954	4.7	64.4	6.8
☐ 7	7	1955	5.5	96.9	9.9
☐ 8	8	1956	6.3	93.9	15.4
☐ 9	9	1957	5.6	129.9	18.7
☐ 10	10	1958	7.2	131.3	19.7
☐ 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

2.10.4 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.



2.10.5 Merge Data Files

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

MyBox Merge Text Data Files

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

Source files Source format Target format Target files Logs

Charset

☒ Determine automatically

☐ Known UTF-8

☒ First line defines columns names

Text delimiter

☐ String

☐ Tab ☐ Blank ☐ Blank characters ☐ 4 blanks ☐ 8 blanks

☒ , ☐ | ☐ # ☐ - ☐ + ☐ : ☐ ; ☐ @ ☐ & ☐ % ☐ ! ☐ " ☐ ' ☐ ? ☐ . ☐ * ☐ \ ☐ / ☐ _ ☐ = ☐ < ☐ >

The first line of the first file determines target file's fields.

2.11 Data in System Clipboard

1. Read and parse contents in system clipboard.
2. Delimiter can be chosen from special characters or inputted regular expression.
3. First row can be set as column names.

The screenshot shows the 'Data in System Clipboard' window. On the left, a text area contains the following data:

```
SepalLengthCm,SepalWidt
5.1,3.5,1.4,0.2,Iris-setosa
4.9,3.0,1.4,0.2,Iris-setosa
4.7,3.2,1.3,0.2,Iris-setosa
4.6,3.1,1.5,0.2,Iris-setosa
5.0,3.6,1.4,0.2,Iris-setosa
5.4,3.9,1.7,0.4,Iris-setosa
4.6,3.4,1.4,0.3,Iris-setosa
```

On the right, a table displays the parsed data. The first row is used as column headers. The table contains 7 rows of data.

Source row number	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
Row1	5.1	3.5	1.4	0.2	Iris-setosa
Row2	4.9	3.0	1.4	0.2	Iris-setosa
Row3	4.7				Iris-setosa
Row4	4.6				Iris-setosa
Row5	5.0				Iris-setosa
Row6	5.4				Iris-setosa
Row7	4.6				Iris-setosa

A dialog box titled 'MyBox: Set of Easy Tools' is open, showing the 'Text delimiter' settings. The 'String' option is selected, and a list of special characters is displayed for selection. The 'First line defines the columns' names checkbox is checked in the main window.

2.12 Data in MyBox Clipboard

The screenshot displays the MyBox Clipboard application interface. The main window is titled "Data in MyBox Clipboard : MyBox Clipboard - 6 - b". The interface is divided into several sections:

- Top Menu Bar:** Includes "Window", "Document", "Image", "Data", "File", "Media", "Network", "Settings", "Recent Accessed", "Development", and "Help".
- Toolbar:** Contains various icons for file operations, editing, and viewing.
- Main Data Table:** A table with columns: "Ta...", "ID", "Type", "Name", "Row...", "Colu...", and "+". It contains 7 rows of data. The 7th row is selected (highlighted in grey).
- Dialog Box:** A small dialog box titled "Data in MyBox Clipboa..." is open. It contains a question mark icon and the text "Current name:b". Below this, there is a text input field labeled "New name" with the value "bm" entered. At the bottom are two buttons: "确定" (Confirm) and "取消" (Cancel).
- Bottom Status Bar:** Shows "Page 1 / 1", "Rows: 7/7 Selected: 1", and "Page size 50".

The right side of the image shows a partial view of another application window, which appears to be a table editor with columns "Tab...", "Data row", "收入" (Income), and "快乐" (Happiness). It also has a similar toolbar and status bar.

2.13 Matrix

2.13.1 Edit and Manage Matrices

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

MyBox Manage Matrices : Matrix - 1065 - a

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

Columns Attributes View Edit

Ta...	ID	Type	Name	Row...	Colu...	+
☒ 1	1065	Matrix	a	3	3	
☐ 2	1064	Matrix	b	6	4	

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

Table

Ta...	Data row	列1	列2	列3	+
☐ 1	1	353.63	469.35	59.21	
☐ 2	2	687.2	344.24	308.21	
☐ 3	3	359.8	359.52	925.57	

Text

2.13.2 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.

MyBox Matrix Unary Calculation : Matrix - 1065 - a

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

☐

☐ Ta...	ID	Type	Name	Row...	Colu... +
☒ 1	1065	Matrix	a	3	3
☐ 2	1064	Matrix	b	6	4

☐

Data Calculation Result

☐ Transpose ☐ Row echelon form ☐ Reduced row echelon form
☐ Complement minor ☒ Normalize ☐ 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 ☐ Rows ☐ All

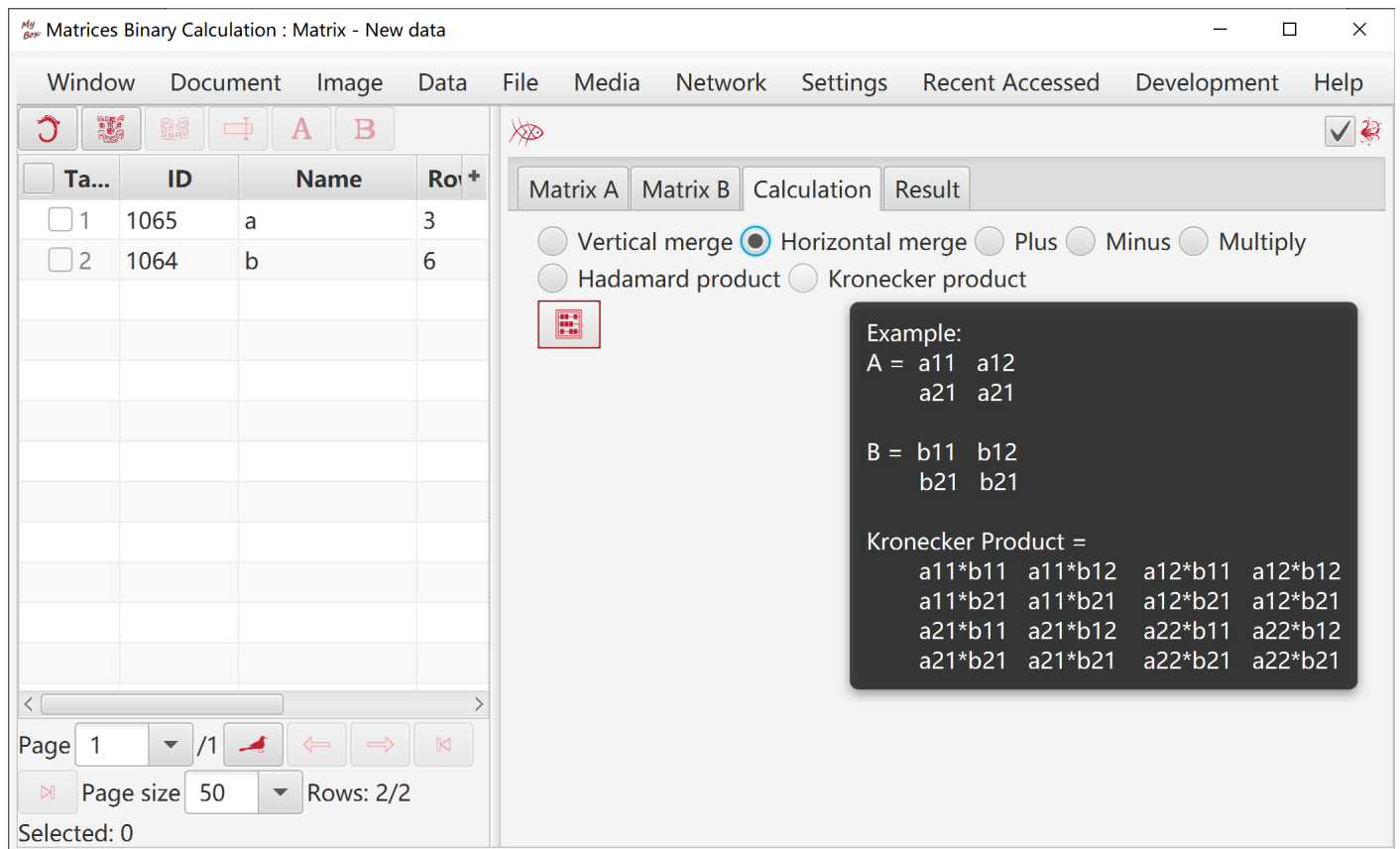
Algorithm ☒ MinMax ☐ L1(Sum) ☐ L2(Z-Score)

Range ,

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

2.13.3 Binary Matrices Calculation

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



2.14 Database Tables

2.14.1 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 V Table definition

Ta...	ID	Type	Name	Row...	Colt +
☒ 1	1113	Databas...	IncomeHappi...	477	3
☐ 2	1081	Databas...	ChinaPopulat...	73	7
☐ 3	1069	Databas...	ed	3	3
☐ 4	1068	Databas...	ChinaPopulat	73	7
☐ 5	9				
☐ 6	4				
☐ 7	3				
☐ 8	9				

Table definition

Ta...	Data row	id	income	happiness
☐ 1	1	1	3.8626474...	2.3144889...

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

50

/10

2.14.2 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. 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.

3. When MyBox create name of table/column:

- Invalid characters are converted as underscore characters.
- If it does not start with a letter, character "a" is added in front of it.

Database Table | Data | Logs

Table name

☒ Create auto-increment column as primary key

☐ Select columns as primay keys

↑ ↓ ⇅

Ta...	Column	
☐ 1	Column1	
☐ 2	Column2	
☐ 3	Column3	

SQL identifiers

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.

2.14.3 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 'Database SQL: 848 - Maximum rows' window. The left sidebar contains a tree view under 'SQL' with 'Examples' expanded, showing 'Query meta of database' and 'Query tables of database'. 'Maximum rows' is selected under 'Query tables of database'. The main area has a text input for SQL statements with the example: 'SELECT * FROM visit_history FETCH FIRST 300 ROWS ONLY'. Below the input is a 'Results' tab showing a table with 8 rows and 7 columns: Data row, RESOURC..., FILE_TYPE, OPERATIO..., RESOURC..., DATA_MO..., and L/. The table contains data for rows 1 through 8. The bottom status bar shows 'Selected: 0', 'Rows: 20/300', 'Page size: 20', 'Page: 1 / 15'.

Database SQL: 848 - Maximum rows

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

Nodes Tags Time Find

Select

- ▼ SQL
 - ▼ Examples
 - ▼ Query meta of database
 - Names of user tables
 - Information of user tables
 - Names of system tables
 - Information of system tables
 - Names of views
 - Information of views
 - Names and types of table
 - Information of table columns
 - Names of table indices
 - Information of table indices
 - ▼ Query tables of database
 - Range of rows
 - Maximum rows**
 - First row
 - Conditions of query
 - Count

Attributes SQL

Input statements of SQL. One statement in each line.

SELECT * FROM visit_history FETCH FIRST 300 ROWS ONLY

Results Data

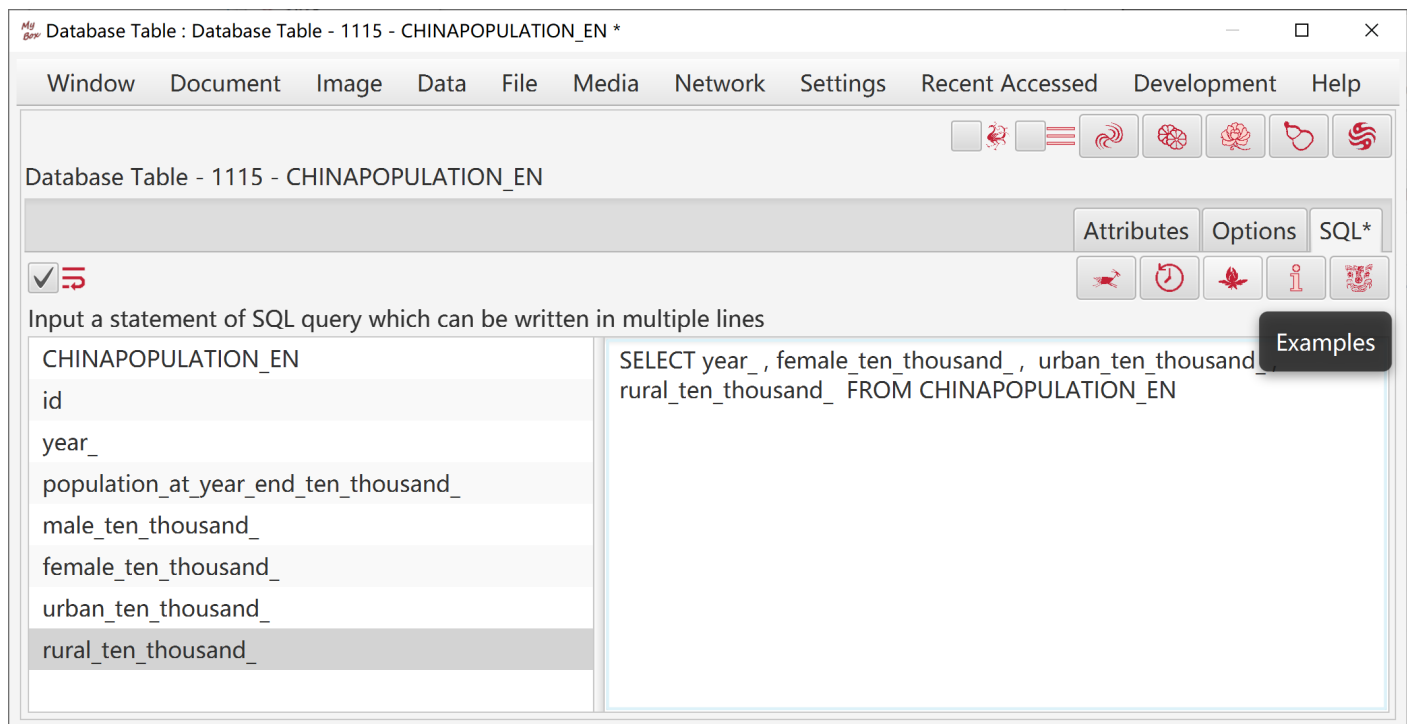
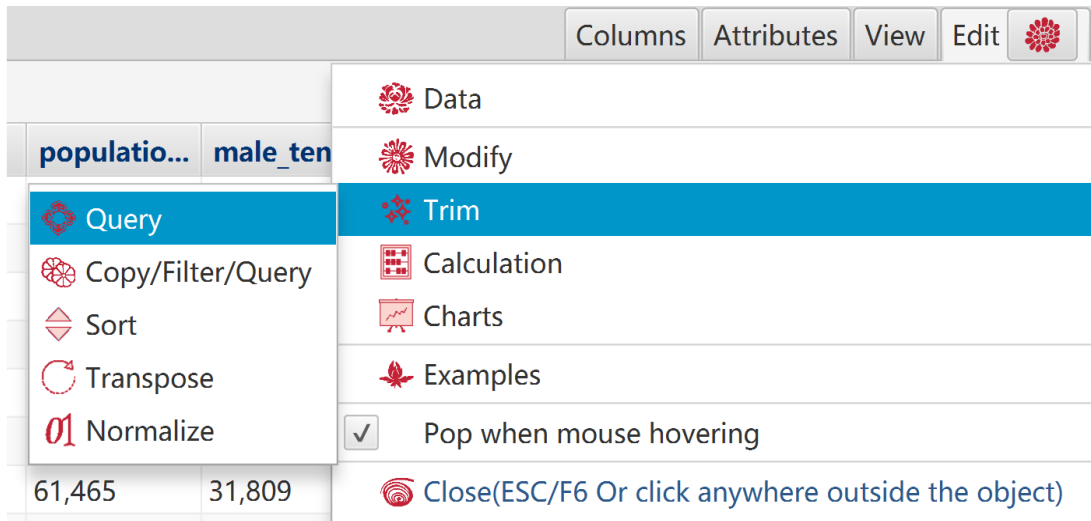
Data row	RESOURC...	FILE_TYPE	OPERATIO...	RESOURC...	DATA_MO...	L/ +
1	2	21	3	d:\tmp\my...		20
2	1	21	3	d:\tmp\my...		20
3	2	21	2	d:\tmp\my...		20
4	1	21	2	d:\tmp\my...		20
5	2	21	3	d:\tmp\my...		20
6	1	21	3	d:\tmp\my...		20
7	2	21	2	d:\tmp\my...		20
8	1	21	2	d:\tmp\my...		20

Selected: 0 Rows: 20/300 Page size 20 Page 1 / 15

2.14.4 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.



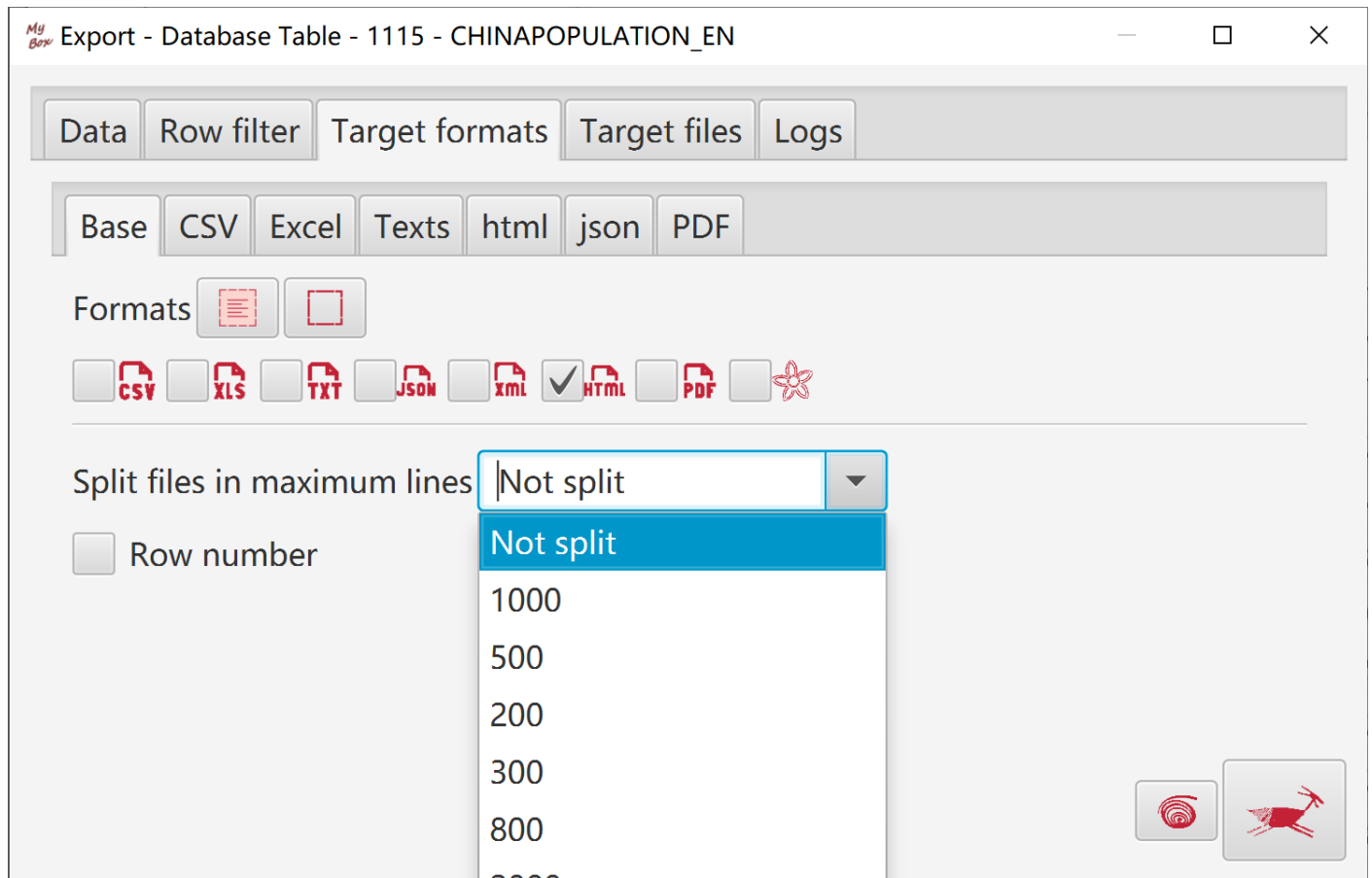
2.15 Export

1. Select 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.

2. Select targets and their formats: csv, texts, excel, xml, json, html, pdf, MyBox Clipboard.

3. Split files in maximum lines.



2.16 Convert to Database Table

1. Select 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.

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

3. Option to import data.

MyBox Convert to database table - Database Table - 1115 - CHINAPOPULATION_EN

Convert to database table

Data Row filter Options Database Table Logs

Columns   No selection means all columns

☐ id ☐ year_ ☐ populationat_year_end_ten_thousand_ ☐ maleten_thousand_ ☐ femaleten_thousand_ ☐ urbanten_thousand_ ☐ ruralten_thousand_

☒ Import data

To invalid number ☒ Set as blank ☐ Set as zero

3 Calculation Tools

3.1 JShell(Java interactive coding tool)

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.

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+I" 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

- ▼ Examples
 - ▼ Object
 - string
 - time
 - array
 - list
 - ▼ Expressions
 - Numeric operation
 - Strings operations
 - ▼ Boolean operation
 - list include
 - string match
 - string include
 - and/or/not
 - ▼ Methods
 - Area of circle
 - Round value
 - Format number**
 - Format date
 - string match
 - string include

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).doubleValue();
}

import java.text.DecimalFormat;
String formatDouble(double data, int scale) {
    try {
        String format = "#,###";
        if (scale > 0) {
            format += "." + "#".repeat(scale);
        }
        DecimalFormat df = new DecimalFormat(format);
        return df.format(data);
    } catch (Exception e) {
        return "";
    }
}

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();
}

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
formatDouble(circleAreaByRadius(273.4), 4)

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

Loaded

3.2 JEXL(Java Expression Language)

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.

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 enrionment 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 paramters have valid values, result is shown in right pane.
7. JEXL codes are organized in tree. Examples are provided.

The screenshot shows the MyBox JEXL editor interface. The title bar indicates the file is 'JEXL(Java Expression Language): 952 - percentage'. The menu bar includes Window, Document, Image, Data, File, Media, Network, Settings, Recent Accessed, Development, and Help. The sidebar on the left, titled 'Select', shows a tree of JEXL Codes. Under 'JEXL Codes', 'Examples' is expanded, showing 'Object', 'Math', and 'Boolean expression'. Under 'Boolean expression', 'percentage' is selected. A tooltip for the 'percentage' function is visible, showing the code: `var value = 2; var set = {1,'2',3, 'hello'}; value !~ set;`. The main editor area has two tabs: 'JEXL script' and 'JEXL Context'. The 'JEXL script' tab is active, showing the following code: `var number = 37; var total = 518; var scale = 2; DoubleTools.percentage(number, total, scale);`. The 'JEXL Context' tab shows the code: `jexlContext.set("DoubleTools", mara.mybox.tools.DoubleTools.class);`. The 'JEXL script paramters' section is empty. The right panel, titled 'Results', shows three execution results for different snippets. Each result includes a timestamp, the snippet code, and a box with details: id, Status, Type, and Name.

JEXL Codes Sidebar:

- ▼ JEXL Codes
 - ▼ Examples
 - Object
 - Math
 - ▼ Boolean expression
 - in array
 - not in set
 - match string
 - include string
 - and/or/not
 - ▼ Numeric operations
 - decimal scale
 - format
 - percentage**
 - ▼ Date
 - now
 - format
 - ▼ Statement
 - loop
 - function

JEXL script:

```
var number = 37;
var total = 518;
var scale = 2;
DoubleTools.percentage(number, total, scale);
```

JEXL Context:

```
jexlContext.set("DoubleTools", mara.mybox.tools.DoubleTools.class);
```

JEXL script paramters:

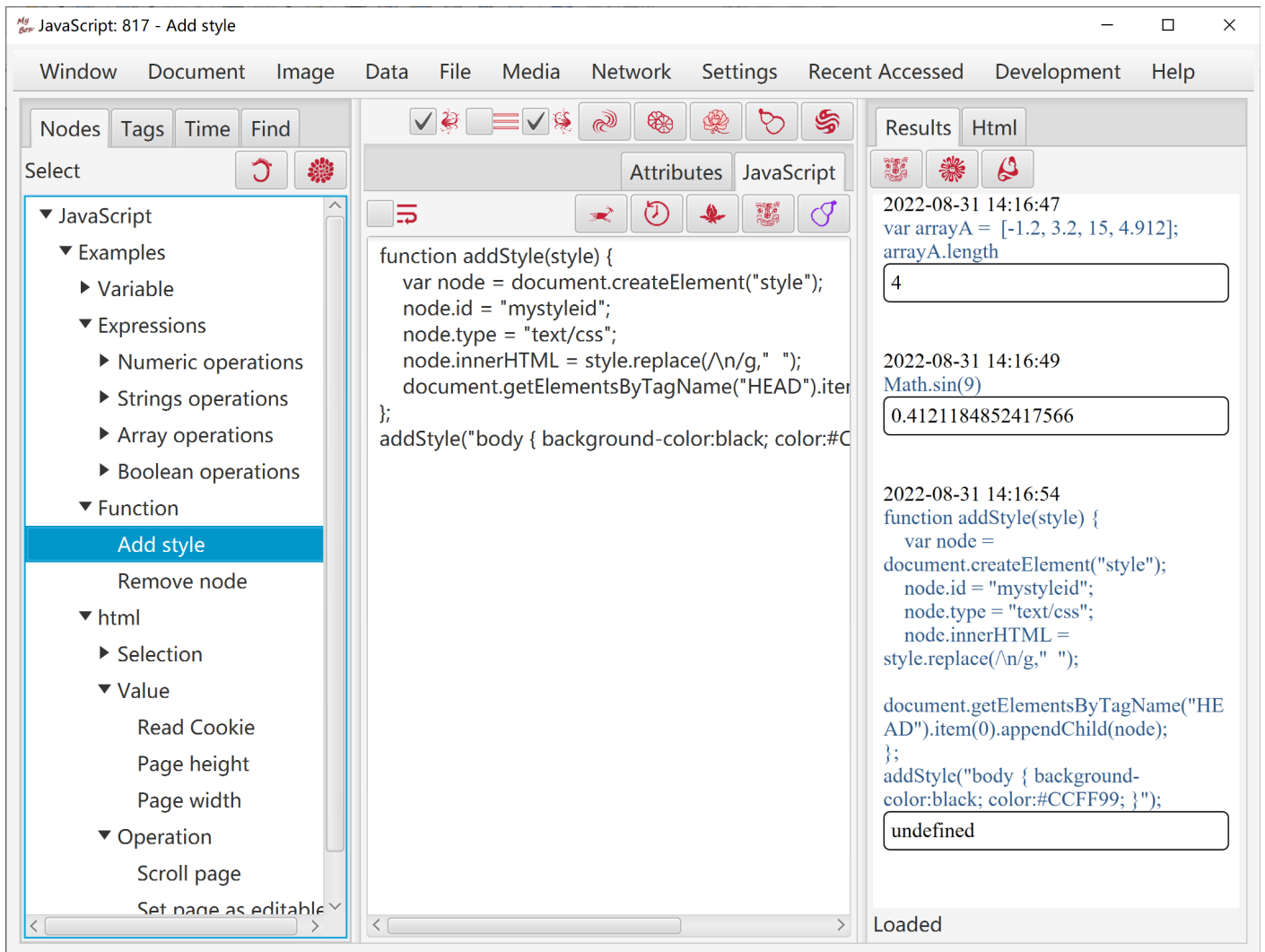
Results:

- 2022-08-31 14:13:54
jexlContext.set("DoubleTools", mara.mybox.tools.DoubleTools.class);
id: 29
Status: Valid
Type: STATEMENT
- 2022-08-31 14:13:54
jexlScript =
jexlEngine.createScript("var number = 37; var total = 518; var scale = 2; DoubleTools.percentage(number, total, scale);");
id: 30
Status: Valid
Type: EXPRESSION
Name: jexlScript
- 2022-08-31 14:13:54
jexlScript.execute(jexlContext);
id: 31
Status: Valid
Type: VAR
Name: \$7

3.3 Javascript

This tool helps to edit, run, save Javascript codes.

Javascript codes are organized in tree. Examples are provided.



4 Data of Location

4.1 Data Constraints

4.1.1 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)

4.1.2 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.

4.1.3 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``.

4.1.4 Time

4.1.4.1 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".

4.1.4.2 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"

4.1.4.3 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

4.2 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 condition.
 - Manage conditions: Add/delete/edit/copy.
 - Conditions ever executed are saved automatically.
 - Recently visited conditions are listed in pop window of the buttons.

4.3 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.

4.4 Geography Code

4.4.1 Data Definition

1. Basic 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.

4.4.2 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.

4.4.3 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.

4.4.4 Define Condition

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

4.4.5 Import Data

4.4.5.1 Predefined Data

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

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

4.4.5.2 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

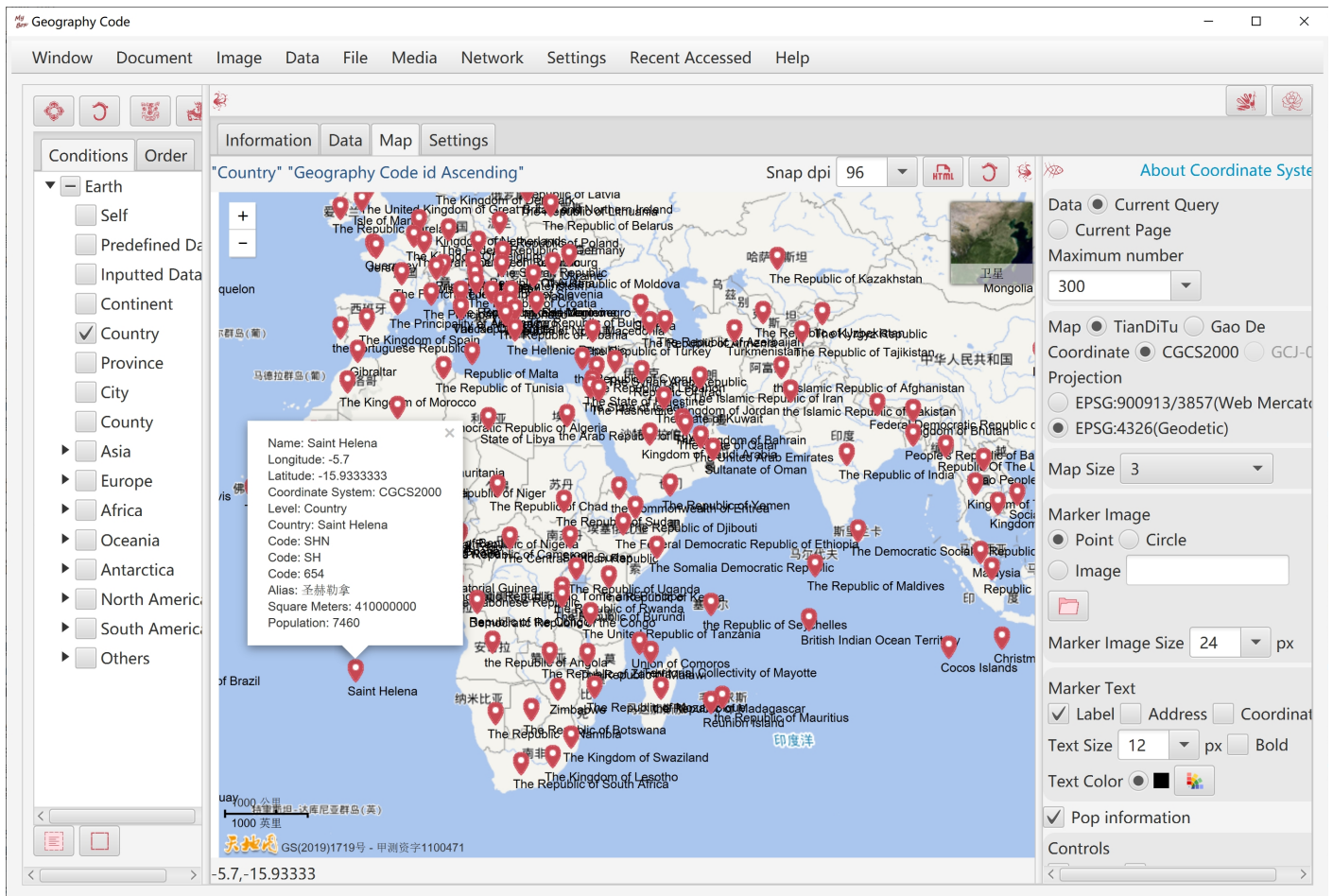
4.4.5.3 Data from geoname.org:

- 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.

4.4.6 Settings

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

MyBox User Guide – Data Tools v6.5.9

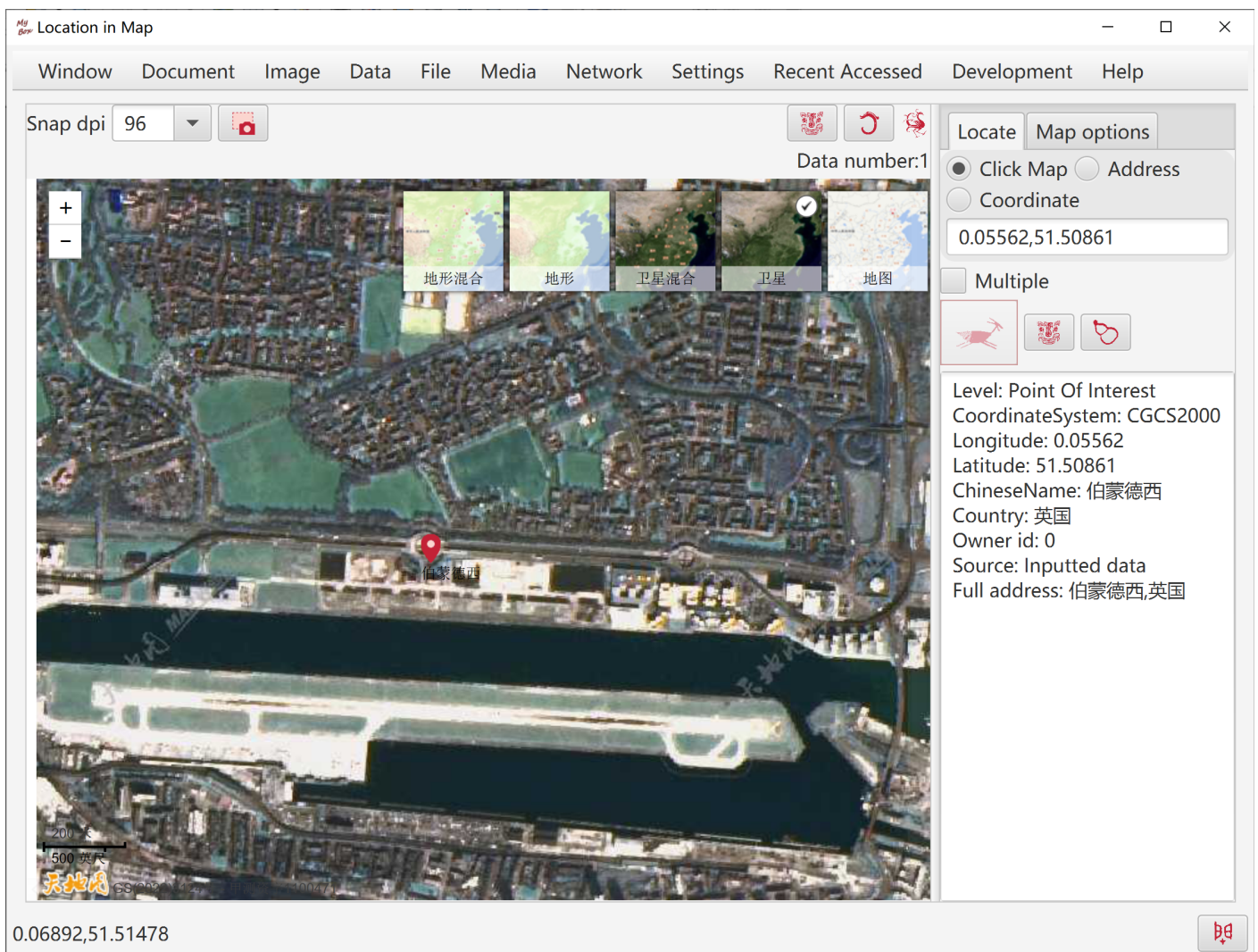


4.5 Location in Map

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.



4.6 Location Data

4.6.1 Data Definition

1. Basic attributes: data set, label, longitude, latitude, start time, end time.
2. Auxiliary attributes: altitude, precision, coordinate system, speed, direction, data value, data size, image, comments.

4.6.2 Data constraints

1. Each location data belongs to a data set.
2. Data set defines common attributes of some location data, examples:
 - Date format
 - Whether omit "AD" for date AD
 - Text color
 - Image

These attributes help to distinguish data points in map.

4.6.3 Define Conditions

1. List of data sets. Multiple nodes can be selected.
2. Time tree(Start time). Multiple nodes can be selected.

4.6.4 Map Data

1. At beginning, the first data is made as map center.
2. Location Distribution: All data are displayed in map.
3. Time Sequence:
 - Data are displayed in frames as "Start Time" in ascending order.
 - When "Accumulated" is selected, drawn points will not be erased and points are shown more and more.
 - When "Time Overlay" is selected, all data whose duration(between "start time" and "end time") has intersection with duration of current frame will be treated as valid points of current frame.
Example, current frame has "start time" of "1044 BC" and "end time" of "221 BC", then all data who or part of who appears in this duration will be displayed in current frame.
 - When "Move Center" is selected, each frame will adjust its map center.
 - When "Link" is selected, lines between adjacent 2 points will be shown.
 - Control frames:
 - Set interval.
 - Select a frame(by start time).
 - Pause/Continue playing.
 - Previous/Next frame.
 - Whether loop.

4.6.5 Snapshots

- For "Location Distribution":
 - html:Data and snapshot of current frame
 - Snapshot of current frame. All supported image formats can be selected.
- For "Time Sequence", more choices:
 - jpg:Snapshots of all frames
 - png:Snapshots of all frames
 - Animated gif:Snapshots of all frames(May out of memory)

4.6.6 Import Data

If data include a dataset which is not in database, the new dataset will be added in database automatically.

4.6.6.1 CSV Format

- Necessary fields: Dataset,Longitude,Latitude
- Optional fields: Label,Address,Altitude,Precision,Speed,Direction,Coordinate System, Data Value,Data Size,Start Time,End Time,Image,Comments

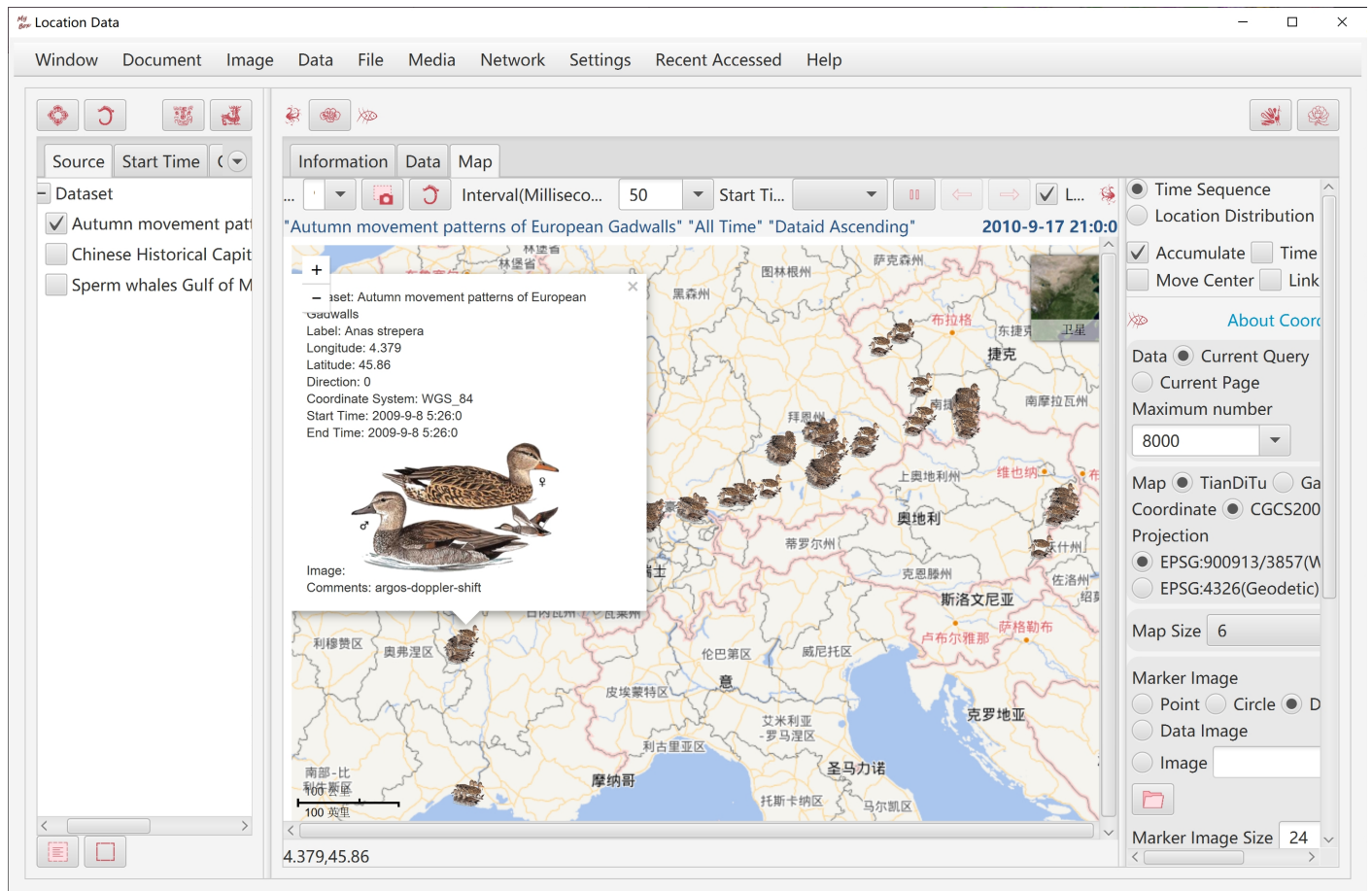
4.6.6.2 Data from movebank.org

- Download address:
 - <https://www.datarepository.movebank.org/>
- Comma-delimited CSV file.
- Necessary fields: timestamp,location-long,location-lat,study-name
- Coordinate system is WGS_84.

4.6.6.3 Examples

1. Chinese Historical Capitals
2. Autumn movement patterns of European Gadwalls
3. Sperm whales Gulf of Mexico

MyBox User Guide – Data Tools v6.5.9



4.7 Location Tools

1. 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

2. Convert coordinate values as other coordinate systems.

Location Tools

Window

Document

Image

Data

File

Media

Network

Settings

Recent Accessed

Help

Degrees

48

Minutes

51

Seconds

12.2796

Degrees Minutes Seconds(DMS)

48°51'12.2796"

Decimal

48.853411

Coordinate System

☒ CGCS2000(China Geodetic Coordinate System)
 ☐ GCJ-02(China encrypted coordinate)
 ☐ WGS-84(GPS)
 ☐ BD-09(Baidu encrypted coordinate)
 ☐ Mapbar coordinate

Longitude

117.0983

Latitude

36.25551

[About Coordinate System](#)

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

4.8 Epidemic Reports

4.8.1 Data Definition

1. Basic attributes: dataSet, time, location, source.
2. Basic values: confirmed, healed, dead.
3. Subtraction statistic: increased confirmed, increased healed, increased dead.
Calculated by adjacent rows.
4. Division statistics:
 - healed/confirmed permillage, dead/confirmed permillage
 - confirmed/population permillage, healed/population permillage, dead/population permillage,
 - confirmed/area permillage, healed/area permillage, dead/area permillage.

When value of "area"/"population" of location is invalid(zero or negative), corresponding statistics data are meaningless.

Predefined data "countries" have valid "area"/"population" and they have meaningful statistics values.
5. Accumulation statistics:
 - Values of some countries. Calculated by values of country's provinces.
 - Values of continents. Calculated by values of continent's countries.
 - Values of Earth. Calculated by values of continents.

4.8.2 Data Constraints

1. Not null values: dataSet, time, location
2. Values of "source": "Inputted data", "Predefined data", "Filled data", "Statistics data".
3. "location" is foreign key of "Geography Code", which must have row defined in that table.
4. In "confirm", "healed", "dead", at least one should be larger than zero.
5. One of following can determine a data row:
 - Match id, which is assigned by MyBox automatically. This is accurate matching.
 - Match "dataSet" + "date" + "location". This is accurate matching.
 - This version assumes that only one valid data in each day for same dataSet plus same location.

4.8.3 Edit Data

1. When in single data, location is set by selecting node from locations tree.
2. In interface of "Epidemic Reports of Chinese Provinces" or "Epidemic Reports of Countries", multiple rows can be inputted for same dataSet and time.
3. Modify values of "source" for selected data rows.

4.8.4 Import Data

4.8.4.1 Predefined Data

COVID-19 historical data from Johns Hopkins University.(Till 2020-09-24)

4.8.4.2 CSV format:

- Download address:
https://github.com/Mararsh/MyBox_data/tree/master/md/COVID19/en
- Necessary fields: Data Set,Time,Confirmed,Healed,Dead,
And location data which are enough to define a geography code:
Longitude,Latitude,Level,Continent,Country,Province,County,Town,Village,Building,Point of Interest
- Optional fields: Increased Confirmed,Increased Healed,Increased Dead
- Coordinate system is CGCS2000.

4.8.4.3 COVID-19 historical data from Johns Hopkins University(Global) :

- Download address:
https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series
- Necessary fields: Province/State,Country/Region,Lat,Long.
And date list like "1/22/20,1/23/20..."
- Coordinate system is WGS_84.
- Australia, Canada and China are reported at the province/state level, and others are at country level.
- Items whose values are all zero will be skipped.

4.8.4.4 COVID-19 daily data from Johns Hopkins University(Global) :

- Download address:
https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_daily_reports
- Data fields are change as time flowing...
Following is format of "01-22-2020.csv":
Province/State,Country/Region,Last Update,Confirmed,Deaths,Recovered
Following is format of "05-15-2020.csv":
FIPS,Admin2, Province_State,Country_Region, Last_Update,Lat,Long_,
Confirmed, Deaths, Recovered,Active, Combined_Key
- Coordinate system is WGS_84.
- Items whose values are all zero will be skipped.

4.8.4.5 Handle Data of Imported

1. Option: Statistics against dataset.

2. Time of all data are changed as “23:59:00”.
3. If data include a geography code which is not in database, the new geography code will be added in database automatically.

4.8.5 Statistics Data

1. Option to accumulate date
2. Option to calculate subtraction statistic for different location levels.

4.8.6 Define Conditions

1. Data Sources Tree: Data sets and their different sources are organized a tree. Multiple nodes can be selected.
2. Locations Tree: All geograhpy codes in MyBox are organized as a tree by their subordination relationship. Multiple nodes can be selected.
3. Times Tree: All times involved in Epidemic Reports of MyBox are organized as a tree. Multiple nodes can be selected.
4. Number of Top Data in Each Day:
 - Unlimit. Charts will not be displayed. Data are queried as condition.
 - Valid value:
 - Data are queried as condition, and then be truncated as top data of each day, by which charts and data are displayed.
 - "Time Descending" is always as the first ordering element automaectically.
 - At least one more column should be picked as ordering element.
 - Beside "Time Descending", the first ordering element is called "major quering attribute".
 - Number of Top Data in Each Day" and "Elements of ordering" work for Query and Export, and not for Clear.
5. Edit condition: title, where, order by, fetch, "Number of Top Data in Each Day"(0 or -1 means Unlimit), which are combined together as the final condition.

4.8.7 Display Charts

1. Only when query condition satisfies requirements, charts are displayed. Charts' data are always "Daily top data" and have "Major query attribute".
2. Beside "Major query attribute" , more attribues can be selected, to display multiple dimension data in same chart, or show multiple charts at same time.
3. Chart type: horizontal bars, vertical bars, horizontal lines, vertical lines, pie, map.
4. When there are multiple times in data, charts are animated. Data charts of each time are displayed frame by frame in time ascending.
5. For animated charts, support Pause/Continue, Jump to frame of a time, Last frame, Next frame, setting interval.
6. Common settings, which take effect immediately:
 - Legend location: not display, top, bottom, left, right.

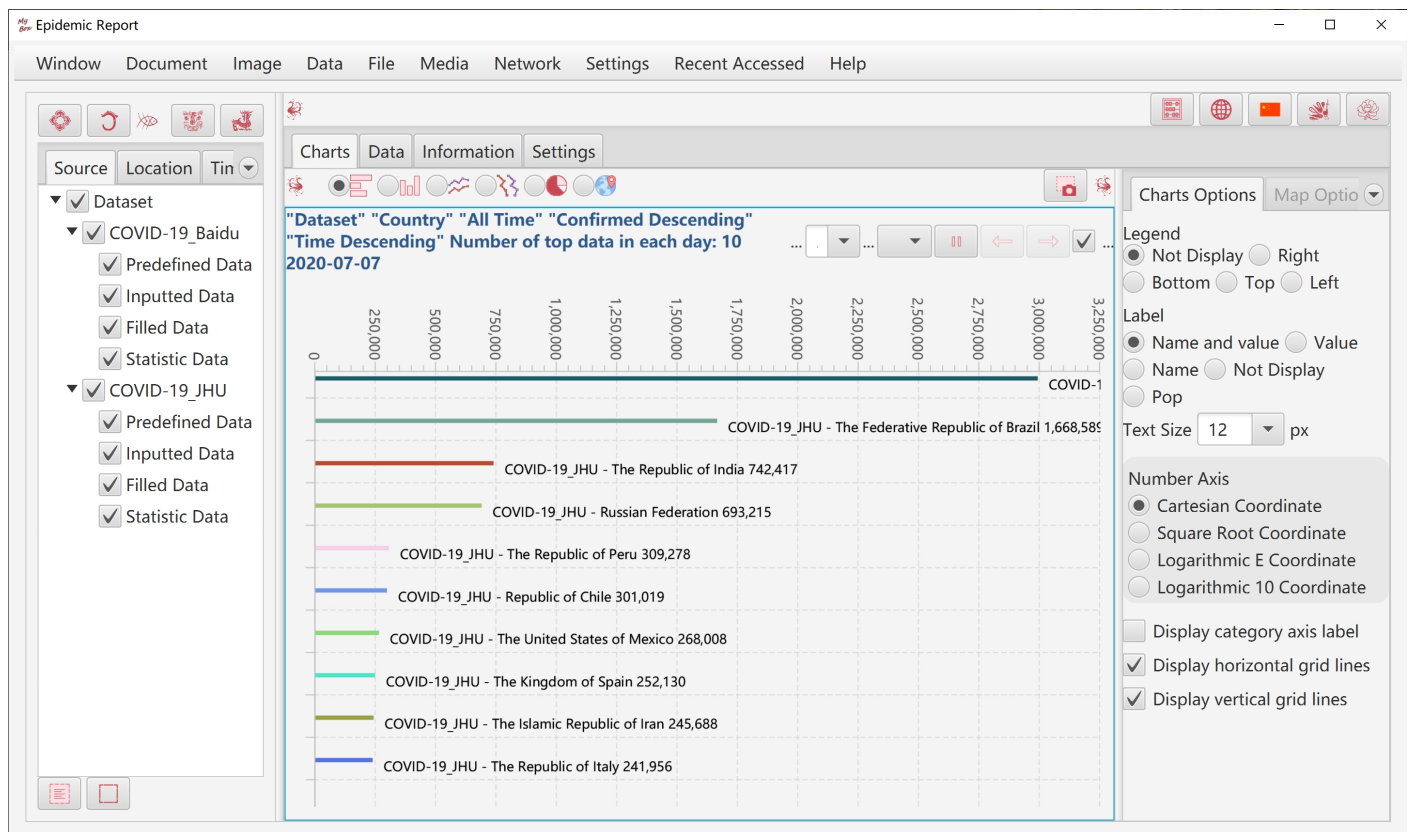
- Value's label: name and value, value, name, not display, pop
- Whether display: category axis, horizontal grid lines, vertical grid lines
- Number axis: Cartesian Coordinate, Square Root Coordinate, Logarithmic E Coordinate, Logarithmic 10 Coordinate.
- Font size
- Parameters of map: level, layers, language

7. Snap chart.

- Snapshot of current frame. All supported image formats can be selected.
- jpg:Snapshots of all frames
- png:Snapshots of all frames
- Animated gif:Snapshots of all frames(May out of memory)

4.8.8 Settings

1. Snap dpi, maximum width of snapped animated images, time to loading chart's data.
These parameters are related to memory usage and computer's calculation capacity.
2. Customize colors of data rows as column "source". Provide "Default" and "Random" buttons.
3. Customize colors of data values in charts. Provide "Default" and "Random" buttons.
4. Customize colors of location values in charts. Provide "Random" buttons.



5 Others

5.1 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.

5.2 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.

5.3 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.

5.4 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.

5.5 Extract ttf files from ttc file

<End of Document>