

## Convert to Binary

**Author: Jesse Dupre, Golden Helix, Inc.**

### Overview

This script prompts the user to enter a threshold value. Based on the threshold value, this script creates a new spreadsheet that converts all real-valued (R) and integer (I) columns from an original spreadsheet to binary (B) columns containing only ones and zeros.

Values  $\geq$  threshold will be converted to 1s and values  $<$  threshold *or missing* will be converted to 0s.

Possible uses include converting haplotype frequencies to binary values, dichotomizing clusters of runs of homozygosity, and converting 1s and 2s (e.g. 1 = control, 2 = case) to 0s and 1s.

### Recommended Directory Location

A good directory location for saving this script at would be:

**C:/HelixTree/scriptsHT/user/Spreadsheet/Edit/.**

When saved to this folder, the script will be accessible from the spreadsheet edit menu.

### Using the Script

Open a spreadsheet containing integer or real-value data types, denoted by an 'I' or 'R' respectively at the top each column. The spreadsheet may also contain other data types.

**Note:** If you do not want every real-valued and integer column to be converted, it is recommended you first create a column subset spreadsheet, run the script to convert the columns, and then join the new spreadsheet with your original spreadsheet.

Unsort	B	1	R	2	R	3	R	4	C
	PATID	BP_I		BP		Age		AGE?	
1	1	0		89.4		41.1		41.1	
2	2	0		101.7		39.6		39.6	
3	3	0		91.1		38.7		38.7	
4	4	1		108.1		61.1		?	
5	5	1		110.3		61		?	
6	6	1		109.6		68.2		?	
7	7	0		96.7		28.7		28.7	
8	8	0		97.2		32.1		32.1	
9	9	0		99.1		56		56	
10	10	0		93.2		34.8		34.8	
11	11	1		111.1		60.4		?	
12	12	0		82.6		41.6		41.6	
13	13	0		90.4		32.6		32.6	
14	14								

Select the **Convert to Binary** script from the spreadsheet menu location where it was stored. Or, from the project viewer, select **Tools->Run Script** to open a dialog to browse for the script location.

A dialog will open asking you to enter a threshold value. Values  $\geq$  threshold will be converted to 1s and values  $<$  threshold *or missing* will be converted to 0s.

**Note:** Missing values (denoted by '?' in the spreadsheet) will be converted to zeros because Binary columns in HelixTree cannot contain missing values. If you do want missing values converted to zeros we recommend you first create a row subset spreadsheet with only those rows not containing ?s.

Enter threshold:

(Values  $\geq$  threshold = 1. Values  $<$  threshold or missing = 0)

OK Cancel

Enter a threshold value and click **OK**.

The script will then create a new spreadsheet with all the integer and real-value columns from the original spreadsheet converted to binary columns containing only ones and zeros.

Unsort		B 1	B 2	B 3	B 4	C 5
	PATID	BP_I	BP	Age	AGE?	Sex
1	1	0	0	0	0	
2	2	0	1	0	0	
3	3	0	1	0	0	
4	4	1	1	0	0	
5	5	1	1	0	0	
6	6	1	1	0	0	
7	7	0	1	0	0	
8	8	0	1	0	0	
9	9	0	1	0	0	
10	10	0	1	0	0	
11	11					