

Column Concordance

Author: Greta Linse, Golden Helix

Overview

This script tests to see if the values of two columns are identical. Output is the original spreadsheet with an additional concordance column. In the concordance column a 1 represents matching values and a 0 represents non-matching values.

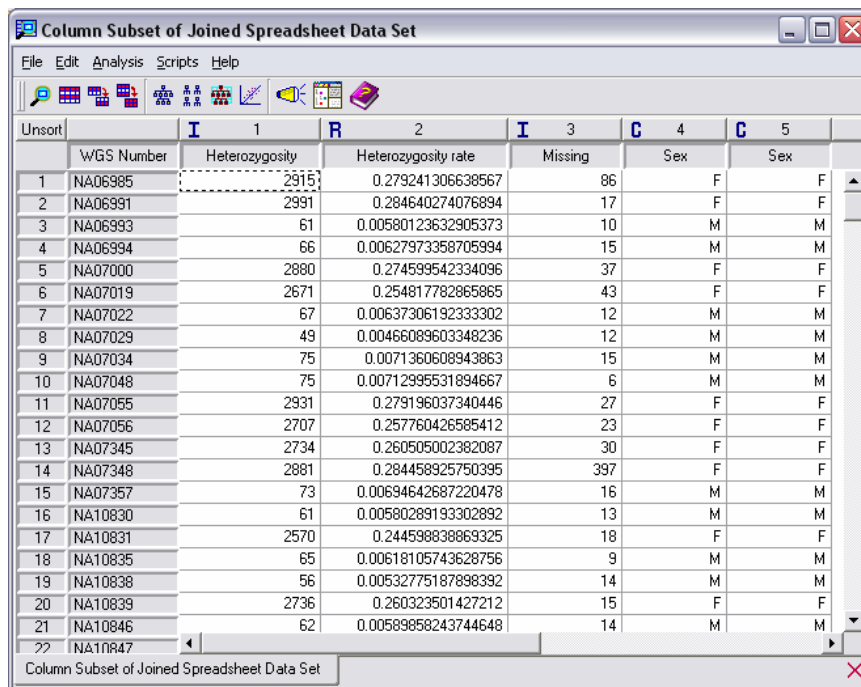
Recommended Directory Location

A good directory location for saving this script would be:
C:/HelixTree/scriptsHT/user/Spreadsheet/Scripts/.

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

Using the Script

To use this script open a spreadsheet with two columns containing the data you want to check concordance on (Figure 1).



Unsort	I	1	R	2	I	3	C	4	C	5
	WGS Number	Heterozygosity	Heterozygosity rate	Missing	Sex	Sex				
1	NA06985	2915	0.279241306638567	86	F	F				
2	NA06991	2991	0.284640274076894	17	F	F				
3	NA06993	61	0.00580123632905373	10	M	M				
4	NA06994	66	0.00627973358705994	15	M	M				
5	NA07000	2880	0.274599542334096	37	F	F				
6	NA07019	2671	0.254817782865865	43	F	F				
7	NA07022	67	0.00637306192333302	12	M	M				
8	NA07029	49	0.00466089603348236	12	M	M				
9	NA07034	75	0.0071360608943863	15	M	M				
10	NA07048	75	0.00712995531894667	6	M	M				
11	NA07055	2931	0.279196037340446	27	F	F				
12	NA07056	2707	0.257760426585412	23	F	F				
13	NA07345	2734	0.260505002382087	30	F	F				
14	NA07348	2881	0.284458925750395	397	F	F				
15	NA07357	73	0.00694642687220478	16	M	M				
16	NA10830	61	0.00580289193302892	13	M	M				
17	NA10831	2570	0.244598838869325	18	F	F				
18	NA10835	65	0.00618105743628756	9	M	M				
19	NA10838	56	0.00532775187898392	14	M	M				
20	NA10839	2736	0.260323501427212	15	F	F				
21	NA10846	62	0.00589858243744648	14	M	M				
??	NA11847									

Figure 1. Spreadsheet.

Select the **Column Concordance** 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.

Running the script will prompt you to enter the numbers of the two columns you want to compare (Figure 2). The above example would be columns 4 and 5 respectively.

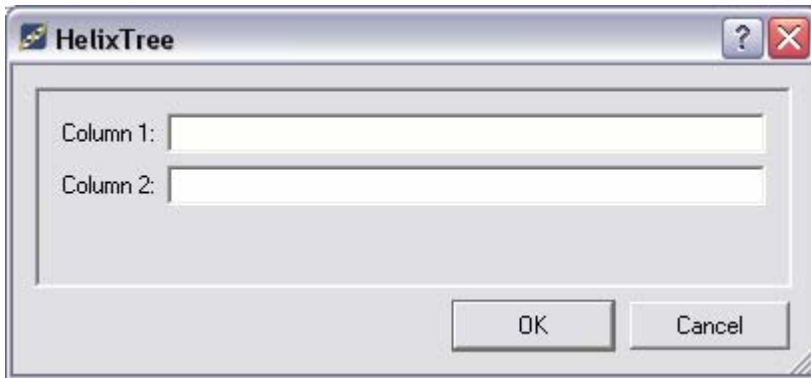
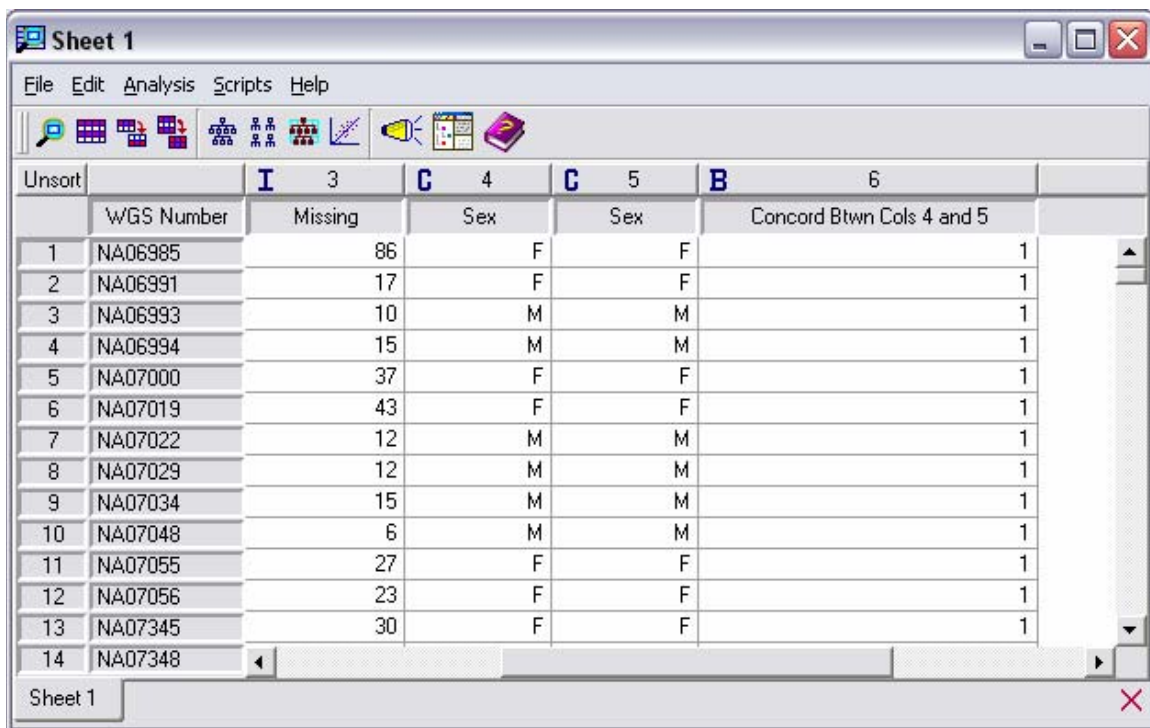


Figure 2. Column Concordance dialog window.

This will produce a new spreadsheet (Figure 3) with an additional concordance column indicating whether or not the values of each column match. 1 represents matching values and 0 non-matching values.

A spreadsheet window titled "Sheet 1" with a menu bar (File, Edit, Analysis, Scripts, Help) and a toolbar. The spreadsheet has columns: Unsort, WGS Number, Missing, Sex (col 4), Sex (col 5), and Concord Btwn Cols 4 and 5. The data rows are numbered 1 to 14. The "Concord Btwn Cols 4 and 5" column contains values 1 for all rows, indicating a match. The "Missing" column contains values for each row, and the "Sex" columns contain 'F' or 'M'.

Unsort	WGS Number	Missing	Sex	Sex	Concord Btwn Cols 4 and 5
1	NA06985	86	F	F	1
2	NA06991	17	F	F	1
3	NA06993	10	M	M	1
4	NA06994	15	M	M	1
5	NA07000	37	F	F	1
6	NA07019	43	F	F	1
7	NA07022	12	M	M	1
8	NA07029	12	M	M	1
9	NA07034	15	M	M	1
10	NA07048	6	M	M	1
11	NA07055	27	F	F	1
12	NA07056	23	F	F	1
13	NA07345	30	F	F	1
14	NA07348				

Figure 3. Gender Concordance output with non-matching values highlighted.