4j – A Guide to
Mann-Whitney U Test

Royal Geographical Society with IBG
Advancing geography and geographical learning

The Mann-Whitney U Test is used to analyse whether two data samples are significantly different from one another or whether any differences witnessed by the researcher are there simply due to chance.

Why would we use the Mann-Whitney U test?

Researchers who are interested in how similar two sets of data are, rather than if there is a correlation between those two sets are best using this test. It can be used when two samples are clearly independent from one another. The data should also be made up of ordinal data (data that can be placed in a clear rank order).

Worked Example:

For this worked example, the following data was gathered, showing how questionnaire participants rated the quality of service provision for two towns, Town A and Town B.
Ratings were given on a 0 to 5 scale.

Participant number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Town A	3	1	2	2	0	2	3	1	0	1	1	2	0	1	3
Town B	3	2	3	4	2	5	3	4	1	4	2	4	4	1	5

1. Place the values of both samples combined in numerical order, noting to which sample (in this case which town) each piece of data refers. If there are two of the same value, place that from Town A first in the list. It does not matter which sample you choose to be first in this case, as long as you are consistent.

A	A	A	A	A	A	A	A	B	B	A	A	A	A	B	B	B	A	A	A	B	B	B	B	B	B	B	B	B	B
0	0	0	1	1	1	1	1	1	1	2	2	2	2	2	2	2	3	3	3	3	3	3	4	4	4	4	4	5	5

2. For each value for Town B, count how many Town A values come before it in the list. Add together these counts to get a U_1 value.

A	A	A	A	A	A	A	A	B	B	A	A	A	A	B	B	B	A	A	A	B	B	B	B	B	B	B	B	B	B
0	0	0	1	1	1	1	1	1	1	2	2	2	2	2	2	2	3	3	3	3	3	3	4	4	4	4	4	5	5
								8	8					12	12	12				15	15	15	15	15	15	15	15	15	15

U_1 = 8 + 8 + 12 + 12 + 12 + 15 + 15 + 15 + 15 + 15 + 15 + 15 + 15 + 15 + 15

U_1 = 202

3. For each value for Town A, count how many Town B values come before it in the list. Add together these counts to get a U_2 value.

A	A	A	A	A	A	A	A	B	B	A	A	A	A	B	B	B	A	A	A	B	B	B	B	B	B	B	B	B	B
0	0	0	1	1	1	1	1	1	1	2	2	2	2	2	2	2	3	3	3	3	3	3	4	4	4	4	4	5	5
0	0	0	0	0	0	0	0			2	2	2	2				5	5	5										

U_2 = 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 2 + 2 + 2 + 2 + 5 + 5 + 5

U_2 = 23

4. Using a critical value table, one can tell if this result is significant or not. One compares the smaller of the U values with the critical value read from the table. The copy below can be easily found online for use with higher n_1 and n_2 values. The table below gives critical values to 5% significance.

	n_2	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
n_1																			
1																			
2									0	0	0	0	1	1	1	1	1	2	2
3						0	1	1	2	2	3	3	4	4	5	5	6	6	7
4					0	1	2	3	4	4	5	6	7	8	9	10	11	11	12
5				0	1	2	3	5	6	7	8	9	11	12	13	14	15	17	18
6				1	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24
7				1	3	5	6	8	10	12	14	16	18	20	22	24	26	28	30
8			0	2	4	6	8	10	13	15	17	19	22	24	26	29	31	34	36
9			0	2	4	7	10	12	15	17	20	23	26	28	31	34	37	39	42
10			0	3	5	8	11	14	17	20	23	26	29	33	36	39	42	45	48
11			0	3	6	9	13	16	19	23	26	30	33	37	40	44	47	51	55
12			1	4	7	11	14	18	22	26	29	33	37	41	45	49	53	57	61
13			1	4	8	12	16	20	24	28	33	37	41	45	50	54	59	63	67
14			1	5	9	13	17	22	26	31	36	40	45	50	55	59	64	67	74
15			1	5	10	14	19	24	29	34	39	44	49	54	59	**64**	70	75	80
16			1	6	11	15	21	26	31	37	42	47	53	59	64	70	75	81	86
17			2	6	11	17	22	28	34	39	45	51	57	63	67	75	81	87	93
18			2	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99

The size of each sample is denoted by n_1 and n_2 (in this case, the sample size for both Town A and Town B is the same). Both n_1 and n_2 are 15. This gives us a critical value of 64.

5. In order for the two samples to be thought of as significantly different from each other (and not due to chance) the smaller U value has to be equal to or less than the critical value from the table.

   In the example, the U value of 23 is less than the critical value of 64. Therefore, we can say with 95% certainty that the questionnaire respondents have rated Town A and Town B significantly differently. The next stage for the geographer is to find a reason why the respondents may have reacted this way.