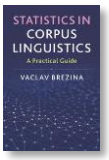


Brezina, V. (2018). *Statistics in Corpus Linguistics: A Practical Guide.* Cambridge University Press.

Group comparison tool instructions

This tool can be used to:

- Compare two or more (sub)corpora.
- Perform parametric (t-test and ANOVA) and non-parametric (Mann-Whitney U test and Kruskal-Wallis) statistical tests.
- Perform a statistical test on different groups or a repeated measures statistical test.

Instructions:

1) Prepare your dataset in Excel or Calc. Include an ID column.

2 groups

ID

	A	B	C
1	ID	Female	Male
2	1	1794.6	1296.1
3	2	1681.3	1402
4	3	1282.5	1427.1
5	4	1435.1	1394.4
6	5	1377.9	1513.7
7	6	1577.8	1529.6
8	7	1432.1	1395.8
9	8	1485.7	1283.9
10	9	1564.5	1413.7
11	10	1609.6	1547.7

3 and more groups

ID

	A	B	C	D	E
1	ID	AB	C1	C2	DE
2	1	1.08	3.85	11.36	16.79
3	2	2.64	2.38	29.73	10.7
4	3	15.06	0	1.06	21.43
5	4	1.3	0	21.16	6.09
6	5	0	1.24	5	0.44
7	6	0.39	2.84	8.77	10.66
8	7	0	1.18	15.69	0
9	8	8.47	15.22	2.26	9.38
10	9	0.52	0	7.97	17.65

2) Copy-paste data in the text-box.

3) Select options

2. Select data options.

◉ Different groups ○ Same group different conditions

3. Select type of test.

◉ Parametric test ○ Non-parametric test

4) Click on 'Compare'.

1. Paste tab delimited data including header row and id column. For help click here.

20	1733.234762	1520.867116
21	1647.409434	1388.107127
22	1721.825963	1249.029955
23	1470.531525	1409.668313
24	1714.000091	1566.951567
25	1443.735563	1750.673965
26	1566.634336	1593.705632
27	1329.434914	1475.874088
28	1848.72324	1166.567637
29	1507.423045	1814.228021
30	1483.620009	1215.845581
31	1389.139198	1527.955193
32	1549.9986	1560.364465

2. Select data options.

◉ Different groups ○ Same group different conditions

3. Select type of test.

◉ Parametric test ○ Non-parametric test

Compare Clear

5) The output

The output shows the results of the chosen statistical test, a relevant effect size measure and 95% confidence interval for this effect size, where available.

degrees of freedom (df)

test statistic

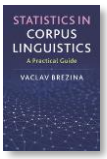
p-value

T-TEST (Welch Two Sample): t (61.93) = 2.77, p = 0.007; Statistically significant? YES
d = 0.69, 95% CI [0.18, 1.21]; MEDIUM EFFECT
r = 0.33, 95% CI [0.08, 0.53]; MEDIUM EFFECT

effect size

95% confidence interval

interpretation of the effect size



One-way ANOVA: $F\,(3, 56) = 5.59$; $p = 0.002$; Statistically significant? YES
$\omega = 0.432$; MEDIUM EFFECT

Post-hoc tests: Bonferroni adjusted t-test - p-values

GROUPS	AB	C1	C2
C1	1.000	-	-
C2	0.015	0.027	-
DE	0.043	0.075	1.000

statistically significant results highlighted

R code that performs the analysis can be viewed and copied when going with the mouse pointer to R code