

Statistics in Microsoft Excel tutorial

Objective

Principal component analysis (PCA), projection to latent structure regression (PLS-R), and projection to latent structure based discriminant analysis (PLS-DA) are the commonly-used multivariate analysis method in metabolomics study. Of course, there are many platforms for these analyses not only by the commercially available software such as SIMCA-P, but also by freely available software such as statistical language R, MetaboAnalyst. I also use them very much! However, I think that excel based platform should be also required for the approachability and the easily adjustment of Figures. In addition to multivariate analysis, I think that the easy-to-use platform for the graph making and the multiple comparison method is also required although I also know these tasks are possible by some tools like VANTED. I would like to offer the platform such that many researchers can easily do the statistical analysis.

Format (Please see the example file)

- ✓ Data file must be saved as comma separated values (CSV) file format.
- ✓ First column must include the labels such as the retention time, peak number, and compound name in addition to two first rows, i.e. "Class", "Row: Label:: Column: Sample" index.
- ✓ First row must include the "Class" index. The order of class number is flexible. But users must use the class number from 1 without skip number.
- ✓ Second row must include the sample information.
- ✓ Others must include variables.

1	Class	B	C	D	E	F	G	H	I	J	K	L	M	N
2	Row Label Column Sample	No1-1	No1-2	No1-3	No6-1	No6-2	No6-3	No11-1	No11-2	No11-3	No16-1	No16-2	No16-3	No21-1
3	2-Hydroxypyrimidine C02502	3735	2722	2391	3435	2147	2407	5118	3753	4402	4245	3599	3588	4383
4	Pyruvate+Oxalacetic acid C00022+C00036	251	155	172	234	163	181	211	194	190	198	181	183	194
5	Lactic acid C00186	557	234	260	425	169	200	251	193	206	151	148	149	148
6	Glycolic acid C00160	323	231	276	182	146	162	215	189	209	205	207	202	131
7	Alanine 2TMS_Major C00041	16355	15239	14263	20288	15808	17903	12653	13660	11157	13821	11732	11832	8748
8	C11_Allane	499	475	482	609	353	418	461	351	406	447	430	366	418
9	m-Butylamine C18706	1390	1336	1456	1242	1157	1282	1207	993	1364	1244	1320	1127	1097
10	Oxalate C00209	108056	79260	83403	19787	14424	14021	2524	25973	26964	43401	30423	31388	15205
11	3-Hydroxybutyrate C01089	681	522	401	725	552	526	658	462	746	1109	1116	887	847
12	2-Aminobutyric acid C00261	190	147	149	181	141	172	95	95	98	154	122	138	75
13	Malonic acid C00383	2008	1389	1341	1349	978	935	1298	1071	952	1499	1182	1078	1146
14	Valine 2TMS_Major C00183	1991	1653	1538	4324	3327	3417	1395	1386	1257	1623	1390	1421	1244
15	Unknown_42	2528	2035	1782	1671	1188	1210	2459	2182	2484	2338	2047	1985	1607
16	Unknown_45	2267	2481	3266	2188	2151	2393	1771	2108	2644	2101	2049	2193	2022
17	Urea C00086	225	169	185	2	2	2	176	153	162	223	206	200	407
18	Serine 2TMS_Minor C00065	5022	4259	4637	3755	3528	4336	3691	2615	4376	2210	3142	2235	3022
19	2-Aminoethanol C00189	1187	1167	1279	1667	1461	1465	725	806	803	1216	1248	1171	1038
20	Unknown_58	20113	18560	18692	20936	16908	19385	10712	10790	11478	16077	16077	15437	9581
21	Phosphate C00009	76942	70625	74834	79784	64091	72110	40373	40944	43067	61382	61136	58624	35877
22	Leucine 2TMS C00123	1187	1291	1216	3885	3635	3443	795	922	772	1359	1182	1191	1097
23	Isoleucine 2TMS C00407	867	918	850	2379	2097	2058	840	922	776	1172	1032	1047	733
24	Proline C00148	2166	2417	2224	3502	3177	2968	1215	1486	1139	1749	1377	1502	1438
25	Malic acid C01384	2	2	3	2	2	2	2	113	209	2	114	176	64
26	Glycine 2TMS_Major C00037	1269	1102	1112	1388	1178	1355	861	804	839	818	785	763	789
27	Succinic acid(or aldehyde) C00042	18422	16495	16589	12101	10329	11123	12149	11815	12142	12776	12599	11686	9682
28	Glycine acid C00028	1719	1528	1612	1344	1156	1216	2355	2584	2584	2881	2854	2729	2034
29	Fumaric acid C00122	728	665	653	704	637	695	591	588	604	617	617	593	546
30	Serine 2TMS_Major C00065	12226	14159	14118	11054	10360	11490	7814	9620	7993	9488	8585	9231	7734
31	Pipecolic acid C00408	357	450	380	166	143	164	186	304	167	218	122	206	121
32	Unknown_85_Sugar like	3138	3067	3659	3256	3185	3043	3434	3427	3729	2776	3103	2791	2617
33	Threonine 2TMS C00188	2764	2949	2796	4672	4535	4906	2208	2508	2247	2829	2652	2841	2109
34	3-Methylglutarate PubChem12284	432	294	321	356	291	336	278	338	355	227	283	273	207
35	Unknown_84	2	129	116	219	221	224	112	116	84	99	70	100	121

Available methods

1. Normalization method by internal standard
2. Bar graph making
3. Line chart making
4. Multi T-test
5. Principal component analysis (PCA)
6. Projection to latent structure regression (PLS-R)
7. Projection to latent structure based discriminant analysis (PLS-DA)

