Distribution of Parameters for Generated SQL Queries on the TPC-H Database

Scenario: 1-6jp_ujjp__g__h0-3ss_h0nss

Scale factor: 5 GB

The 1000-query pack was generated on: 2023-11-11

First column ('Query parameter or value') displays either:

- a parameter name, or

- a value of the current parameter (only for low-cardinality parameters).

Second column ('Descriptive statistics') displays:

- for low cardinality parameters: frequency and proportion from total (of 1000 queries)

- for higher cardinality parameters: [min, median, max] mean / SD

Query parameter or value	Descriptive statistics
SELECT_n_of_columns	[1, 9, 189] 15 / 23
SELECT_n_of_non_aggr_func___LOWER	[0.00, 0.00, 12.00] 0.61 / 1.12
SELECT_n_of_non_aggr_func___RTRIM	[0.00, 0.00, 9.00] 0.65 / 1.15
SELECT_n_of_non_aggr_func___UPPER	[0.00, 0.00, 10.00] 0.61 / 1.09
SELECT_n_of_non_aggr_func___SQRT	
0	796 (80%)
1	140 (14%)
2	27 (2.7%)
3	15 (1.5%)
4	8 (0.8%)
5	4 (0.4%)
6	7 (0.7%)
7	2 (0.2%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___SUBSTR	[0.00, 0.00, 11.00] 0.66 / 1.14

Query parameter or value	Descriptive statistics
SELECT_n_of_non_aggr_func___LTRIM	
0	590 (59%)
1	271 (27%)
2	84 (8.4%)
3	22 (2.2%)
4	17 (1.7%)
5	7 (0.7%)
6	3 (0.3%)
7	3 (0.3%)
8	3 (0.3%)
SELECT_n_of_non_aggr_func___TRUNC	
0	901 (90%)
1	77 (7.7%)
2	20 (2.0%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func___DOW	
0	847 (85%)
1	110 (11%)
2	22 (2.2%)
3	12 (1.2%)
4	3 (0.3%)
5	4 (0.4%)
6	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___MONTH	
0	810 (81%)
1	128 (13%)
2	24 (2.4%)
3	21 (2.1%)
4	6 (0.6%)
5	4 (0.4%)
6	3 (0.3%)
8	4 (0.4%)
SELECT_n_of_non_aggr_func___YEAR	
0	827 (83%)
1	118 (12%)
2	37 (3.7%)
3	7 (0.7%)
4	4 (0.4%)
5	3 (0.3%)
6	1 (0.1%)
7	2 (0.2%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func___FLOOR	
0	917 (92%)
1	64 (6.4%)

Query parameter or value	Descriptive statistics
2	15 (1.5%)
3	2 (0.2%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func___DAY	
0	878 (88%)
1	84 (8.4%)
2	24 (2.4%)
3	7 (0.7%)
4	5 (0.5%)
5	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___ROUND	
0	914 (91%)
1	69 (6.9%)
2	14 (1.4%)
3	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___LOG	
0	845 (85%)
1	117 (12%)
2	19 (1.9%)
3	12 (1.2%)
4	2 (0.2%)
5	3 (0.3%)
6	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___ABS	
0	910 (91%)
1	71 (7.1%)
2	13 (1.3%)
3	6 (0.6%)
SELECT_n_of_all_non_aggr_func	[0, 3, 79] 5 / 9
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 13.00] 0.89 / 2.05
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 13.00] 0.87 / 1.96
SELECT_n_of_aggr_func___AVG	
0	934 (93%)
1	48 (4.8%)
2	16 (1.6%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 13.00] 0.81 / 1.88
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 12.00] 0.87 / 1.89
SELECT_n_of_aggr_func___SUM	
0	933 (93%)
1	42 (4.2%)
2	16 (1.6%)

Query parameter or value	Descriptive statistics
3	4 (0.4%)
4	4 (0.4%)
5	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.0, 3.0, 14.0] 3.6 / 3.2
FROM_n_of_join_paths	
1	160 (16%)
2	170 (17%)
3	159 (16%)
4	150 (15%)
5	165 (17%)
6	196 (20%)
FROM_n_of_super_joins___LEFT	
0	552 (55%)
1	303 (30%)
2	117 (12%)
3	24 (2.4%)
4	4 (0.4%)
FROM_n_of_super_joins___RIGHT	
0	537 (54%)
1	313 (31%)
2	126 (13%)
3	22 (2.2%)
4	1 (0.1%)
5	1 (0.1%)
FROM_n_of_super_joins___FULL	
0	542 (54%)
1	310 (31%)
2	109 (11%)
3	35 (3.5%)
4	4 (0.4%)
FROM_n_of_joins___INNER	[0.00, 3.00, 12.00] 3.35 / 2.61
FROM_n_of_joins___RIGHT	[0.00, 3.00, 13.00] 3.38 / 2.50
FROM_n_of_processed_rows	[5, 47,300,128, 226,499,509] 61,486,496 / 49,173,476
WHERE_n_of_predicates	[0.00, 3.00, 13.00] 3.78 / 2.98
WHERE_n_of_attribs_of_type___character_varying	
0	378 (38%)
1	281 (28%)
2	181 (18%)
3	92 (9.2%)
4	43 (4.3%)
5	20 (2.0%)
6	4 (0.4%)
7	1 (0.1%)
WHERE_n_of_attribs_of_type___integer	
0	374 (37%)
1	292 (29%)

Query parameter or value	Descriptive statistics
2	191 (19%)
3	89 (8.9%)
4	37 (3.7%)
5	11 (1.1%)
6	5 (0.5%)
8	1 (0.1%)
WHERE_n_of_attribs_of_type___numeric	
0	586 (59%)
1	262 (26%)
2	97 (9.7%)
3	43 (4.3%)
4	11 (1.1%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type___character	
0	655 (66%)
1	245 (25%)
2	79 (7.9%)
3	18 (1.8%)
4	2 (0.2%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type___date	
0	781 (78%)
1	177 (18%)
2	38 (3.8%)
3	3 (0.3%)
5	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	532 (53%)
1	305 (31%)
2	114 (11%)
3	41 (4.1%)
4	6 (0.6%)
5	2 (0.2%)
WHERE_n_of_connect_OR	[0.00, 1.00, 12.00] 2.16 / 2.48
WHERE_n_of_operators___greater_or_less	
0	334 (33%)
1	264 (26%)
2	209 (21%)
3	102 (10%)
4	50 (5.0%)
5	30 (3.0%)
6	8 (0.8%)
7	1 (0.1%)
8	2 (0.2%)
WHERE_n_of_operators___in	
0	531 (53%)

Query parameter or value	Descriptive statistics
1	301 (30%)
2	113 (11%)
3	35 (3.5%)
4	19 (1.9%)
5	1 (0.1%)
WHERE_n_of_operators___between	
0	560 (56%)
1	265 (27%)
2	122 (12%)
3	37 (3.7%)
4	14 (1.4%)
5	2 (0.2%)
WHERE_n_of_operators___like	
0	781 (78%)
1	187 (19%)
2	29 (2.9%)
3	2 (0.2%)
4	1 (0.1%)
WHERE_n_of_non_aggr_func___SQRT	
0	875 (88%)
1	113 (11%)
2	12 (1.2%)
WHERE_n_of_non_aggr_func___DAY	
0	966 (97%)
1	34 (3.4%)
WHERE_n_of_non_aggr_func___FLOOR	
0	957 (96%)
1	42 (4.2%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___ROUND	
0	954 (95%)
1	45 (4.5%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___YEAR	
0	962 (96%)
1	38 (3.8%)
WHERE_n_of_non_aggr_func___TRUNC	
0	956 (96%)
1	44 (4.4%)
WHERE_n_of_non_aggr_func___MONTH	
0	963 (96%)
1	36 (3.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___LOG	
0	924 (92%)
1	69 (6.9%)

Query parameter or value	Descriptive statistics
2	7 (0.7%)
WHERE_n_of_non_aggr_func___DOW	
0	952 (95%)
1	48 (4.8%)
WHERE_n_of_non_aggr_func___ABS	
0	961 (96%)
1	38 (3.8%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	607 (61%)
1	280 (28%)
2	75 (7.5%)
3	30 (3.0%)
4	7 (0.7%)
5	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 12.00] 3.09 / 2.96
HAVING_n_of_main_predicates	
0	311 (31%)
1	95 (9.5%)
2	128 (13%)
3	183 (18%)
4	130 (13%)
5	97 (9.7%)
6	56 (5.6%)
HAVING_n_of_main_predicates___scalar_subquery	
0	436 (44%)
1	195 (20%)
2	179 (18%)
3	190 (19%)
HAVING_n_of_subqueries___scalar_subquery	[0.0, 1.0, 23.0] 2.3 / 3.6
HAVING_n_of_processed_rows_by_subqueries	[0, 10,000,323, 2,554,204,260] 139,407,891 / 289,010,770
ORDER_BY_n_of_columns	
0	310 (31%)
1	329 (33%)
2	239 (24%)
3	85 (8.5%)
4	22 (2.2%)
5	7 (0.7%)
6	5 (0.5%)
7	1 (0.1%)
8	2 (0.2%)
limit	[3, 514, 1,000] 506 / 286
offset	[0, 0, 1,000] 255 / 334