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

Scenario: 1-6jp_rjjp__g__h0-3ss_h0-3nss

Scale factor: 100 GB

The 1000-query pack was generated on: 2023-11-12

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, 8, 182] 14 / 22
SELECT_n_of_non_aggr_func___LOG	
0	864 (86%)
1	115 (12%)
2	13 (1.3%)
3	5 (0.5%)
4	2 (0.2%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___LOWER	
0	611 (61%)
1	277 (28%)
11	1 (0.1%)
2	74 (7.4%)
3	24 (2.4%)
4	7 (0.7%)

Query parameter or value	Descriptive statistics
5	3 (0.3%)
6	3 (0.3%)
SELECT_n_of_non_aggr_func___LTRIM	[0.00, 0.00, 12.00] 0.60 / 1.08
SELECT_n_of_non_aggr_func___RTRIM	
0	627 (63%)
1	268 (27%)
2	71 (7.1%)
3	19 (1.9%)
4	8 (0.8%)
5	1 (0.1%)
6	2 (0.2%)
7	2 (0.2%)
8	2 (0.2%)
SELECT_n_of_non_aggr_func___UPPER	[0.00, 0.00, 12.00] 0.59 / 1.10
SELECT_n_of_non_aggr_func___DAY	
0	877 (88%)
1	99 (9.9%)
2	14 (1.4%)
3	5 (0.5%)
4	2 (0.2%)
5	2 (0.2%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___MONTH	
0	799 (80%)
1	150 (15%)
2	20 (2.0%)
3	12 (1.2%)
4	9 (0.9%)
5	4 (0.4%)
6	3 (0.3%)
7	2 (0.2%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	629 (63%)
1	272 (27%)
10	1 (0.1%)
2	65 (6.5%)
3	20 (2.0%)
4	9 (0.9%)
5	1 (0.1%)
6	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___DOW	
0	839 (84%)
1	120 (12%)
2	28 (2.8%)

Query parameter or value	Descriptive statistics
3	9 (0.9%)
4	1 (0.1%)
5	2 (0.2%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___ABS	
0	927 (93%)
1	68 (6.8%)
2	3 (0.3%)
3	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___YEAR	
0	842 (84%)
1	118 (12%)
2	22 (2.2%)
3	3 (0.3%)
4	10 (1.0%)
5	2 (0.2%)
6	3 (0.3%)
SELECT_n_of_non_aggr_func___FLOOR	
0	934 (93%)
1	57 (5.7%)
2	5 (0.5%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___SQRT	
0	814 (81%)
1	135 (14%)
2	33 (3.3%)
3	8 (0.8%)
4	7 (0.7%)
5	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___ROUND	
0	917 (92%)
1	75 (7.5%)
2	6 (0.6%)
3	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___TRUNC	
0	935 (94%)
1	57 (5.7%)
2	5 (0.5%)
3	2 (0.2%)
4	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 76.0] 4.6 / 7.1
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 13.00] 0.84 / 1.96

Query parameter or value	Descriptive statistics
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 12.00] 0.80 / 1.83
SELECT_n_of_aggr_func___AVG	
0	918 (92%)
1	57 (5.7%)
2	17 (1.7%)
3	8 (0.8%)
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 13.00] 0.95 / 2.06
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 14.00] 0.92 / 2.02
SELECT_n_of_aggr_func___SUM	
0	941 (94%)
1	41 (4.1%)
2	15 (1.5%)
3	3 (0.3%)
SELECT_n_of_all_aggr_func	[0.00, 3.00, 16.00] 3.71 / 3.07
FROM_n_of_join_paths	
1	164 (16%)
2	172 (17%)
3	175 (18%)
4	168 (17%)
5	152 (15%)
6	169 (17%)
FROM_n_of_super_joins___FULL	
0	540 (54%)
1	312 (31%)
2	125 (13%)
3	21 (2.1%)
4	2 (0.2%)
FROM_n_of_super_joins___LEFT	
0	546 (55%)
1	300 (30%)
2	124 (12%)
3	28 (2.8%)
4	2 (0.2%)
FROM_n_of_super_joins___RIGHT	
0	550 (55%)
1	322 (32%)
2	103 (10%)
3	23 (2.3%)
4	2 (0.2%)
FROM_n_of_joins___INNER	[0.00, 3.00, 11.00] 3.26 / 2.44
FROM_n_of_joins___RIGHT	[0.00, 3.00, 14.00] 3.31 / 2.51
FROM_n_of_processed_rows	[5, 930,037,930, 4,495,227,462] 1,167,392,752 / 961,332,298
WHERE_n_of_predicates	[0.00, 3.00, 13.00] 3.88 / 3.02
WHERE_n_of_attribs_of_type___character	
0	644 (64%)

Query parameter or value	Descriptive statistics
1	251 (25%)
2	83 (8.3%)
3	20 (2.0%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type___character_varying	
0	363 (36%)
1	283 (28%)
2	184 (18%)
3	79 (7.9%)
4	61 (6.1%)
5	23 (2.3%)
6	6 (0.6%)
7	1 (0.1%)
WHERE_n_of_attribs_of_type___integer	
0	382 (38%)
1	273 (27%)
2	187 (19%)
3	91 (9.1%)
4	50 (5.0%)
5	15 (1.5%)
6	1 (0.1%)
7	1 (0.1%)
WHERE_n_of_attribs_of_type___numeric	
0	567 (57%)
1	275 (28%)
2	118 (12%)
3	31 (3.1%)
4	9 (0.9%)
WHERE_n_of_attribs_of_type___date	
0	784 (78%)
1	179 (18%)
2	33 (3.3%)
3	4 (0.4%)
WHERE_n_of_pkey_attribs	
0	544 (54%)
1	304 (30%)
2	114 (11%)
3	28 (2.8%)
4	8 (0.8%)
5	2 (0.2%)
WHERE_n_of_connect_OR	[0.00, 2.00, 12.00] 2.27 / 2.47
WHERE_n_of_operators___greater_or_less	
0	356 (36%)
1	228 (23%)
2	207 (21%)
3	111 (11%)

Query parameter or value	Descriptive statistics
4	57 (5.7%)
5	22 (2.2%)
6	14 (1.4%)
7	3 (0.3%)
8	2 (0.2%)
WHERE_n_of_operators___in	
0	523 (52%)
1	293 (29%)
2	124 (12%)
3	46 (4.6%)
4	11 (1.1%)
5	3 (0.3%)
WHERE_n_of_operators___like	
0	767 (77%)
1	190 (19%)
2	39 (3.9%)
3	3 (0.3%)
4	1 (0.1%)
WHERE_n_of_operators___between	
0	533 (53%)
1	279 (28%)
2	134 (13%)
3	36 (3.6%)
4	15 (1.5%)
5	3 (0.3%)
WHERE_n_of_non_aggr_func___SQRT	
0	890 (89%)
1	102 (10%)
2	5 (0.5%)
3	3 (0.3%)
WHERE_n_of_non_aggr_func___TRUNC	
0	955 (96%)
1	43 (4.3%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___ROUND	
0	957 (96%)
1	42 (4.2%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___DAY	
0	964 (96%)
1	33 (3.3%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func___FLOOR	
0	949 (95%)
1	50 (5.0%)
2	1 (0.1%)

Query parameter or value	Descriptive statistics
WHERE_n_of_non_aggr_func___YEAR	
0	960 (96%)
1	39 (3.9%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___MONTH	
0	973 (97%)
1	27 (2.7%)
WHERE_n_of_non_aggr_func___ABS	
0	948 (95%)
1	50 (5.0%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___DOW	
0	962 (96%)
1	36 (3.6%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___LOG	
0	930 (93%)
1	65 (6.5%)
2	4 (0.4%)
3	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	615 (62%)
1	268 (27%)
2	87 (8.7%)
3	22 (2.2%)
4	7 (0.7%)
5	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 3.00, 12.00] 3.21 / 2.85
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.49 / 2.57
HAVING_n_of_main_predicates___non_scalar_subquery	
0	430 (43%)
1	191 (19%)
2	194 (19%)
3	185 (19%)
HAVING_n_of_main_predicates___scalar_subquery	
0	415 (42%)
1	176 (18%)
2	197 (20%)
3	212 (21%)
HAVING_n_of_subqueries___non_scalar_subquery	
0	430 (43%)
1	191 (19%)
2	194 (19%)
3	185 (19%)
HAVING_n_of_subqueries___scalar_subquery	[0.0, 1.0, 19.0] 2.4 / 3.4

Query parameter or value	Descriptive statistics
HAVING_n_of_processed_rows_by_subqueries	[0, 1,500,075,804, 55,161,956,746] 4,116,816,193 / 6,801,772,643
ORDER_BY_n_of_columns	
0	317 (32%)
1	311 (31%)
2	231 (23%)
3	106 (11%)
4	20 (2.0%)
5	9 (0.9%)
6	5 (0.5%)
7	1 (0.1%)
limit	[6, 520, 999] 516 / 286
offset	[0, 0, 998] 252 / 324