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

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

Scale factor: 2 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, 9, 202] 15 / 22
SELECT_n_of_non_aggr_func___ABS	
0	918 (92%)
1	59 (5.9%)
2	19 (1.9%)
3	4 (0.4%)
SELECT_n_of_non_aggr_func___LOG	
0	878 (88%)
1	87 (8.7%)
2	24 (2.4%)
3	6 (0.6%)
4	4 (0.4%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___RTRIM	
0	631 (63%)

Query parameter or value	Descriptive statistics
1	262 (26%)
2	64 (6.4%)
3	22 (2.2%)
4	10 (1.0%)
5	7 (0.7%)
6	2 (0.2%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	642 (64%)
1	235 (24%)
2	74 (7.4%)
3	28 (2.8%)
4	9 (0.9%)
5	8 (0.8%)
6	2 (0.2%)
7	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___LOWER	
0	611 (61%)
1	266 (27%)
2	77 (7.7%)
3	20 (2.0%)
4	15 (1.5%)
5	4 (0.4%)
6	4 (0.4%)
7	3 (0.3%)
SELECT_n_of_non_aggr_func___UPPER	
0	601 (60%)
1	265 (27%)
2	85 (8.5%)
3	29 (2.9%)
4	8 (0.8%)
5	7 (0.7%)
6	2 (0.2%)
7	3 (0.3%)
SELECT_n_of_non_aggr_func___MONTH	
0	800 (80%)
1	143 (14%)
10	1 (0.1%)
2	28 (2.8%)
3	13 (1.3%)
4	7 (0.7%)
5	3 (0.3%)
7	4 (0.4%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___SQRT	

Query parameter or value	Descriptive statistics
0	801 (80%)
1	154 (15%)
2	19 (1.9%)
3	10 (1.0%)
4	11 (1.1%)
5	2 (0.2%)
6	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___TRUNC	
0	917 (92%)
1	67 (6.7%)
2	11 (1.1%)
3	5 (0.5%)
SELECT_n_of_non_aggr_func___YEAR	
0	824 (82%)
1	124 (12%)
2	31 (3.1%)
3	12 (1.2%)
4	3 (0.3%)
5	3 (0.3%)
6	1 (0.1%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func___DAY	
0	873 (87%)
1	97 (9.7%)
2	23 (2.3%)
3	4 (0.4%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___DOW	
0	859 (86%)
1	100 (10%)
2	29 (2.9%)
3	7 (0.7%)
4	2 (0.2%)
5	1 (0.1%)
6	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___LTRIM	
0	639 (64%)
1	250 (25%)
2	62 (6.2%)
3	28 (2.8%)
4	9 (0.9%)
5	6 (0.6%)
6	2 (0.2%)

Query parameter or value	Descriptive statistics
7	2 (0.2%)
8	2 (0.2%)
SELECT_n_of_non_aggr_func___ROUND	
0	912 (91%)
1	76 (7.6%)
2	10 (1.0%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___FLOOR	
0	917 (92%)
1	61 (6.1%)
2	18 (1.8%)
3	4 (0.4%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 63.0] 4.8 / 7.2
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 14.00] 0.88 / 2.08
SELECT_n_of_aggr_func___SUM	
0	916 (92%)
1	58 (5.8%)
2	14 (1.4%)
3	9 (0.9%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 12.00] 0.87 / 2.02
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 13.00] 0.85 / 1.88
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 11.00] 0.80 / 1.88
SELECT_n_of_aggr_func___AVG	
0	931 (93%)
1	41 (4.1%)
2	20 (2.0%)
3	8 (0.8%)
SELECT_n_of_all_aggr_func	[0.0, 3.0, 15.0] 3.6 / 3.2
FROM_n_of_join_paths	
1	158 (16%)
2	173 (17%)
3	169 (17%)
4	163 (16%)
5	168 (17%)
6	169 (17%)
FROM_n_of_super_joins___RIGHT	
0	543 (54%)
1	320 (32%)
2	112 (11%)
3	24 (2.4%)
4	1 (0.1%)
FROM_n_of_super_joins___FULL	
0	525 (53%)
1	324 (32%)

Query parameter or value	Descriptive statistics
2	113 (11%)
3	32 (3.2%)
4	6 (0.6%)
FROM_n_of_super_joins___LEFT	
0	552 (55%)
1	305 (31%)
2	119 (12%)
3	22 (2.2%)
4	2 (0.2%)
FROM_n_of_joins___INNER	[0.00, 3.00, 13.00] 3.30 / 2.41
FROM_n_of_joins___RIGHT	[0.00, 3.00, 12.00] 3.16 / 2.39
FROM_n_of_processed_rows	[5, 18,515,494, 89,494,471] 23,292,890 / 18,904,914
WHERE_n_of_predicates	[0.00, 3.00, 12.00] 3.93 / 2.98
WHERE_n_of_attribs_of_type___character_varying	
0	362 (36%)
1	276 (28%)
2	187 (19%)
3	113 (11%)
4	45 (4.5%)
5	12 (1.2%)
6	3 (0.3%)
7	2 (0.2%)
WHERE_n_of_attribs_of_type___integer	
0	366 (37%)
1	296 (30%)
2	173 (17%)
3	103 (10%)
4	38 (3.8%)
5	19 (1.9%)
6	5 (0.5%)
WHERE_n_of_attribs_of_type___date	
0	792 (79%)
1	167 (17%)
2	38 (3.8%)
3	2 (0.2%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type___numeric	
0	547 (55%)
1	277 (28%)
2	122 (12%)
3	42 (4.2%)
4	7 (0.7%)
5	5 (0.5%)
WHERE_n_of_attribs_of_type___character	
0	648 (65%)
1	246 (25%)

Query parameter or value	Descriptive statistics
2	80 (8.0%)
3	21 (2.1%)
4	5 (0.5%)
WHERE_n_of_pkey_attribs	
0	553 (55%)
1	272 (27%)
2	126 (13%)
3	38 (3.8%)
4	9 (0.9%)
5	2 (0.2%)
WHERE_n_of_connect_OR	[0.00, 1.00, 11.00] 2.27 / 2.48
WHERE_n_of_operators___between	
0	533 (53%)
1	284 (28%)
2	122 (12%)
3	50 (5.0%)
4	9 (0.9%)
5	2 (0.2%)
WHERE_n_of_operators___greater_or_less	
0	316 (32%)
1	284 (28%)
2	168 (17%)
3	126 (13%)
4	62 (6.2%)
5	31 (3.1%)
6	8 (0.8%)
7	2 (0.2%)
8	3 (0.3%)
WHERE_n_of_operators___like	
0	781 (78%)
1	179 (18%)
2	37 (3.7%)
3	3 (0.3%)
WHERE_n_of_operators___in	
0	532 (53%)
1	295 (30%)
2	118 (12%)
3	38 (3.8%)
4	12 (1.2%)
5	4 (0.4%)
6	1 (0.1%)
WHERE_n_of_non_aggr_func___SQRT	
0	866 (87%)
1	123 (12%)
2	11 (1.1%)
WHERE_n_of_non_aggr_func___DOW	

Query parameter or value	Descriptive statistics
0	962 (96%)
1	38 (3.8%)
WHERE_n_of_non_aggr_func___ROUND	
0	951 (95%)
1	48 (4.8%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___DAY	
0	963 (96%)
1	36 (3.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___ABS	
0	957 (96%)
1	43 (4.3%)
WHERE_n_of_non_aggr_func___MONTH	
0	963 (96%)
1	35 (3.5%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___TRUNC	
0	964 (96%)
1	35 (3.5%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___YEAR	
0	970 (97%)
1	30 (3.0%)
WHERE_n_of_non_aggr_func___LOG	
0	921 (92%)
1	72 (7.2%)
2	7 (0.7%)
WHERE_n_of_non_aggr_func___FLOOR	
0	953 (95%)
1	43 (4.3%)
2	4 (0.4%)
WHERE_n_of_all_non_aggr_func	
0	597 (60%)
1	285 (29%)
2	85 (8.5%)
3	30 (3.0%)
4	3 (0.3%)
GROUP_BY_n_of_columns	[0.00, 2.00, 12.00] 3.13 / 2.89
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.38 / 2.58
HAVING_n_of_main_predicates___non_scalar_subquery	
0	438 (44%)
1	186 (19%)
2	180 (18%)
3	196 (20%)
HAVING_n_of_main_predicates___scalar_subquery	

Query parameter or value	Descriptive statistics
0	439 (44%)
1	185 (19%)
2	183 (18%)
3	193 (19%)
HAVING_n_of_subqueries___non_scalar_subquery	
0	438 (44%)
1	186 (19%)
2	180 (18%)
3	196 (20%)
HAVING_n_of_subqueries___scalar_subquery	[0.0, 1.0, 28.0] 2.5 / 3.7
HAVING_n_of_processed_rows_by_subqueries	[0, 30,541,473, 981,383,634] 88,172,620 / 144,102,374
ORDER_BY_n_of_columns	
0	315 (32%)
1	292 (29%)
2	232 (23%)
3	118 (12%)
4	24 (2.4%)
5	11 (1.1%)
6	3 (0.3%)
7	5 (0.5%)
limit	[2, 507, 1,000] 503 / 286
offset	[0, 0, 1,000] 249 / 327