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

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

Scale factor: 50 GB

The 1000-query pack was generated on: 2023-11-06

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, 6, 91] 10 / 13
SELECT_n_of_non_aggr_func___LOWER	
0	717 (72%)
1	225 (23%)
2	47 (4.7%)
3	8 (0.8%)
4	1 (0.1%)
6	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___RTRIM	
0	705 (71%)
1	229 (23%)
2	52 (5.2%)
3	7 (0.7%)
4	6 (0.6%)

Query parameter or value	Descriptive statistics
5	1 (0.1%)
SELECT_n_of_non_aggr_func___DAY	
0	913 (91%)
1	67 (6.7%)
2	13 (1.3%)
3	5 (0.5%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___DOW	
0	890 (89%)
1	95 (9.5%)
2	10 (1.0%)
3	3 (0.3%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___LTRIM	
0	729 (73%)
1	205 (21%)
2	52 (5.2%)
3	8 (0.8%)
4	3 (0.3%)
5	1 (0.1%)
6	2 (0.2%)
SELECT_n_of_non_aggr_func___MONTH	
0	862 (86%)
1	101 (10%)
2	20 (2.0%)
3	9 (0.9%)
4	4 (0.4%)
5	2 (0.2%)
6	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___UPPER	
0	698 (70%)
1	230 (23%)
2	57 (5.7%)
3	12 (1.2%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___YEAR	
0	871 (87%)
1	103 (10%)
2	16 (1.6%)
3	6 (0.6%)
4	2 (0.2%)
5	1 (0.1%)
6	1 (0.1%)

Query parameter or value	Descriptive statistics
SELECT_n_of_non_aggr_func___SQRT	
0	848 (85%)
1	118 (12%)
2	20 (2.0%)
3	7 (0.7%)
4	6 (0.6%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___LOG	
0	902 (90%)
1	81 (8.1%)
2	12 (1.2%)
3	4 (0.4%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	703 (70%)
1	230 (23%)
2	50 (5.0%)
3	10 (1.0%)
4	5 (0.5%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___FLOOR	
0	937 (94%)
1	59 (5.9%)
2	4 (0.4%)
SELECT_n_of_non_aggr_func___TRUNC	
0	947 (95%)
1	48 (4.8%)
2	3 (0.3%)
3	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___ROUND	
0	951 (95%)
1	45 (4.5%)
2	2 (0.2%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___ABS	
0	953 (95%)
1	43 (4.3%)
2	4 (0.4%)
SELECT_n_of_all_non_aggr_func	[0.00, 2.00, 43.00] 3.08 / 4.33
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 10.00] 0.65 / 1.45
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 9.00] 0.66 / 1.39
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 11.00] 0.60 / 1.49
SELECT_n_of_aggr_func___COUNT_DISTINCT	
0	768 (77%)
1	86 (8.6%)

Query parameter or value	Descriptive statistics
2	54 (5.4%)
3	36 (3.6%)
4	21 (2.1%)
5	16 (1.6%)
6	7 (0.7%)
7	9 (0.9%)
8	3 (0.3%)
SELECT_n_of_aggr_func___SUM	
0	951 (95%)
1	37 (3.7%)
2	8 (0.8%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_aggr_func___AVG	
0	952 (95%)
1	36 (3.6%)
2	7 (0.7%)
3	4 (0.4%)
4	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 11.00] 2.63 / 2.27
FROM_n_of_join_paths	
1	372 (37%)
2	299 (30%)
3	329 (33%)
FROM_n_of_super_joins___RIGHT	
0	784 (78%)
1	201 (20%)
2	15 (1.5%)
FROM_n_of_super_joins___LEFT	
0	773 (77%)
1	196 (20%)
2	31 (3.1%)
FROM_n_of_super_joins___FULL	
0	781 (78%)
1	192 (19%)
2	27 (2.7%)
FROM_n_of_joins___RIGHT	
0	233 (23%)
1	280 (28%)
2	204 (20%)
3	144 (14%)
4	84 (8.4%)
5	34 (3.4%)
6	13 (1.3%)
7	7 (0.7%)
8	1 (0.1%)

Query parameter or value	Descriptive statistics
FROM_n_of_joins___INNER	
0	265 (27%)
1	288 (29%)
2	210 (21%)
3	126 (13%)
4	60 (6.0%)
5	35 (3.5%)
6	13 (1.3%)
7	3 (0.3%)
FROM_n_of_processed_rows	[5, 300,001,594, 1,140,049,247] 293,645,922 / 288,329,553
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.74 / 2.17
WHERE_n_of_attribs_of_type___character_varying	
0	431 (43%)
1	330 (33%)
2	154 (15%)
3	61 (6.1%)
4	20 (2.0%)
5	4 (0.4%)
WHERE_n_of_attribs_of_type___integer	
0	451 (45%)
1	317 (32%)
2	157 (16%)
3	52 (5.2%)
4	18 (1.8%)
5	4 (0.4%)
6	1 (0.1%)
WHERE_n_of_attribs_of_type___numeric	
0	651 (65%)
1	259 (26%)
2	73 (7.3%)
3	14 (1.4%)
4	3 (0.3%)
WHERE_n_of_attribs_of_type___character	
0	746 (75%)
1	202 (20%)
2	42 (4.2%)
3	9 (0.9%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type___date	
0	864 (86%)
1	115 (12%)
2	20 (2.0%)
3	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	615 (62%)

Query parameter or value	Descriptive statistics
1	289 (29%)
2	77 (7.7%)
3	19 (1.9%)
WHERE_n_of_connect_OR	
0	453 (45%)
1	195 (20%)
2	130 (13%)
3	105 (11%)
4	56 (5.6%)
5	37 (3.7%)
6	15 (1.5%)
7	6 (0.6%)
8	3 (0.3%)
WHERE_n_of_operators___between	
0	608 (61%)
1	277 (28%)
2	84 (8.4%)
3	26 (2.6%)
4	5 (0.5%)
WHERE_n_of_operators___greater_or_less	
0	428 (43%)
1	305 (31%)
2	164 (16%)
3	69 (6.9%)
4	28 (2.8%)
5	4 (0.4%)
6	2 (0.2%)
WHERE_n_of_operators___like	
0	818 (82%)
1	165 (17%)
2	15 (1.5%)
3	2 (0.2%)
WHERE_n_of_operators___in	
0	638 (64%)
1	260 (26%)
2	81 (8.1%)
3	14 (1.4%)
4	5 (0.5%)
5	2 (0.2%)
WHERE_n_of_non_aggr_func___TRUNC	
0	976 (98%)
1	23 (2.3%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___DOW	
0	973 (97%)
1	26 (2.6%)

Query parameter or value	Descriptive statistics
2	1 (0.1%)
WHERE_n_of_non_aggr_func___YEAR	
0	980 (98%)
1	20 (2.0%)
WHERE_n_of_non_aggr_func___DAY	
0	980 (98%)
1	20 (2.0%)
WHERE_n_of_non_aggr_func___MONTH	
0	981 (98%)
1	19 (1.9%)
WHERE_n_of_non_aggr_func___LOG	
0	937 (94%)
1	61 (6.1%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___SQRT	
0	912 (91%)
1	82 (8.2%)
2	6 (0.6%)
WHERE_n_of_non_aggr_func___FLOOR	
0	967 (97%)
1	30 (3.0%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func___ABS	
0	970 (97%)
1	29 (2.9%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func___ROUND	
0	971 (97%)
1	29 (2.9%)
WHERE_n_of_all_non_aggr_func	
0	720 (72%)
1	207 (21%)
2	62 (6.2%)
3	8 (0.8%)
4	2 (0.2%)
5	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.29 / 2.05
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.38 / 2.58
HAVING_n_of_main_predicates___non_scalar_subquery	
0	447 (45%)
1	171 (17%)
2	201 (20%)
3	181 (18%)
HAVING_n_of_main_predicates___scalar_subquery	
0	434 (43%)
1	209 (21%)

Query parameter or value	Descriptive statistics
2	183 (18%)
3	174 (17%)
HAVING_n_of_subqueries___non_scalar_subquery	
0	447 (45%)
1	171 (17%)
2	201 (20%)
3	181 (18%)
HAVING_n_of_subqueries___scalar_subquery	[0.0, 1.0, 22.0] 2.3 / 3.5
HAVING_n_of_processed_rows_by_subqueries	[0, 161,461,409, 15,960,689,458] 1,021,065,512 / 1,870,005,324
ORDER_BY_n_of_columns	
0	293 (29%)
1	379 (38%)
2	243 (24%)
3	62 (6.2%)
4	12 (1.2%)
5	7 (0.7%)
6	4 (0.4%)
limit	[2, 488, 1,000] 497 / 288
offset	[0, 36, 998] 260 / 326