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

Scenario: 1-6jp_ujjp__g__h0ss_h0nss

Scale factor: 0.01 GB

The 1000-query pack was generated on: 2023-11-10

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, 163] 16 / 24
SELECT_n_of_non_aggr_func___SUBSTR	
0	617 (62%)
1	262 (26%)
10	1 (0.1%)
2	73 (7.3%)
3	20 (2.0%)
4	10 (1.0%)
5	9 (0.9%)
6	4 (0.4%)
7	4 (0.4%)
SELECT_n_of_non_aggr_func___LTRIM	[0.00, 0.00, 10.00] 0.55 / 0.99
SELECT_n_of_non_aggr_func___UPPER	
0	650 (65%)
1	242 (24%)

Query parameter or value	Descriptive statistics
2	61 (6.1%)
3	22 (2.2%)
4	7 (0.7%)
5	10 (1.0%)
6	5 (0.5%)
7	3 (0.3%)
SELECT_n_of_non_aggr_func___ROUND	
0	925 (93%)
1	61 (6.1%)
2	11 (1.1%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func___RTRIM	[0.00, 0.00, 11.00] 0.62 / 1.16
SELECT_n_of_non_aggr_func___SQRT	
0	794 (79%)
1	159 (16%)
10	1 (0.1%)
2	18 (1.8%)
3	13 (1.3%)
4	7 (0.7%)
5	3 (0.3%)
6	5 (0.5%)
SELECT_n_of_non_aggr_func___LOWER	
0	635 (64%)
1	244 (24%)
2	80 (8.0%)
3	16 (1.6%)
4	14 (1.4%)
5	4 (0.4%)
6	4 (0.4%)
7	1 (0.1%)
8	2 (0.2%)
SELECT_n_of_non_aggr_func___DAY	
0	886 (89%)
1	76 (7.6%)
2	21 (2.1%)
3	11 (1.1%)
4	5 (0.5%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___DOW	
0	848 (85%)
1	103 (10%)
2	26 (2.6%)
3	12 (1.2%)
4	5 (0.5%)
5	5 (0.5%)
6	1 (0.1%)

Query parameter or value	Descriptive statistics
SELECT_n_of_non_aggr_func___MONTH	
0	801 (80%)
1	136 (14%)
2	22 (2.2%)
3	18 (1.8%)
4	8 (0.8%)
5	8 (0.8%)
6	2 (0.2%)
7	4 (0.4%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func___YEAR	
0	829 (83%)
1	118 (12%)
2	27 (2.7%)
3	10 (1.0%)
4	5 (0.5%)
5	5 (0.5%)
6	2 (0.2%)
7	3 (0.3%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___TRUNC	
0	915 (92%)
1	74 (7.4%)
2	9 (0.9%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___LOG	
0	872 (87%)
1	92 (9.2%)
2	19 (1.9%)
3	8 (0.8%)
4	5 (0.5%)
5	2 (0.2%)
6	2 (0.2%)
SELECT_n_of_non_aggr_func___FLOOR	
0	928 (93%)
1	57 (5.7%)
2	10 (1.0%)
3	3 (0.3%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func___ABS	
0	914 (91%)
1	73 (7.3%)
2	11 (1.1%)
3	2 (0.2%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 58.0] 4.9 / 8.2

Query parameter or value	Descriptive statistics
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 13.00] 0.82 / 1.88
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 13.00] 0.95 / 2.14
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 11.00] 0.87 / 1.87
SELECT_n_of_aggr_func___SUM	
0	922 (92%)
1	51 (5.1%)
2	21 (2.1%)
3	5 (0.5%)
4	1 (0.1%)
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 12.00] 0.79 / 1.88
SELECT_n_of_aggr_func___AVG	
0	942 (94%)
1	39 (3.9%)
2	12 (1.2%)
3	6 (0.6%)
4	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.00, 3.00, 15.00] 3.63 / 3.11
FROM_n_of_join_paths	
1	168 (17%)
2	162 (16%)
3	171 (17%)
4	161 (16%)
5	172 (17%)
6	166 (17%)
FROM_n_of_super_joins___LEFT	
0	543 (54%)
1	298 (30%)
2	133 (13%)
3	23 (2.3%)
4	3 (0.3%)
FROM_n_of_super_joins___RIGHT	
0	529 (53%)
1	329 (33%)
2	110 (11%)
3	31 (3.1%)
4	1 (0.1%)
FROM_n_of_super_joins___FULL	
0	546 (55%)
1	316 (32%)
2	112 (11%)
3	22 (2.2%)
4	4 (0.4%)
FROM_n_of_joins___INNER	[0.00, 3.00, 11.00] 3.21 / 2.37
FROM_n_of_joins___RIGHT	[0.00, 3.00, 11.00] 3.30 / 2.45
FROM_n_of_processed_rows	[5, 94,765, 441,850] 119,861 / 94,470
WHERE_n_of_predicates	[0.00, 3.50, 12.00] 3.95 / 2.99

Query parameter or value	Descriptive statistics
WHERE_n_of_attribs_of_type___character_varying	
0	357 (36%)
1	275 (28%)
2	182 (18%)
3	106 (11%)
4	59 (5.9%)
5	13 (1.3%)
6	8 (0.8%)
WHERE_n_of_attribs_of_type___integer	
0	381 (38%)
1	265 (27%)
2	195 (20%)
3	104 (10%)
4	45 (4.5%)
5	8 (0.8%)
6	2 (0.2%)
WHERE_n_of_attribs_of_type___character	
0	650 (65%)
1	252 (25%)
2	77 (7.7%)
3	17 (1.7%)
4	4 (0.4%)
WHERE_n_of_attribs_of_type___numeric	
0	546 (55%)
1	277 (28%)
2	122 (12%)
3	39 (3.9%)
4	15 (1.5%)
6	1 (0.1%)
WHERE_n_of_attribs_of_type___date	
0	782 (78%)
1	174 (17%)
2	35 (3.5%)
3	9 (0.9%)
WHERE_n_of_pkey_attribs	
0	551 (55%)
1	281 (28%)
2	133 (13%)
3	32 (3.2%)
4	2 (0.2%)
5	1 (0.1%)
WHERE_n_of_connect_OR	[0.00, 2.00, 11.00] 2.29 / 2.41
WHERE_n_of_operators___between	
0	509 (51%)
1	307 (31%)
2	125 (13%)

Query parameter or value	Descriptive statistics
3	47 (4.7%)
4	12 (1.2%)
WHERE_n_of_operators___greater_or_less	
0	321 (32%)
1	263 (26%)
2	183 (18%)
3	116 (12%)
4	70 (7.0%)
5	28 (2.8%)
6	14 (1.4%)
7	5 (0.5%)
WHERE_n_of_operators___in	
0	539 (54%)
1	275 (28%)
2	123 (12%)
3	46 (4.6%)
4	14 (1.4%)
5	2 (0.2%)
6	1 (0.1%)
WHERE_n_of_operators___like	
0	788 (79%)
1	186 (19%)
2	21 (2.1%)
3	5 (0.5%)
WHERE_n_of_non_aggr_func___ABS	
0	954 (95%)
1	46 (4.6%)
WHERE_n_of_non_aggr_func___MONTH	
0	958 (96%)
1	41 (4.1%)
2	1 (0.1%)
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	946 (95%)
1	52 (5.2%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___LOG	
0	917 (92%)
1	81 (8.1%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___DAY	
0	969 (97%)
1	30 (3.0%)

Query parameter or value	Descriptive statistics
2	1 (0.1%)
WHERE_n_of_non_aggr_func___YEAR	
0	963 (96%)
1	33 (3.3%)
2	4 (0.4%)
WHERE_n_of_non_aggr_func___SQRT	
0	887 (89%)
1	99 (9.9%)
2	13 (1.3%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func___TRUNC	
0	955 (96%)
1	42 (4.2%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func___DOW	
0	968 (97%)
1	30 (3.0%)
2	2 (0.2%)
WHERE_n_of_all_non_aggr_func	
0	617 (62%)
1	255 (26%)
2	91 (9.1%)
3	29 (2.9%)
4	7 (0.7%)
5	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 13.00] 2.99 / 2.75
HAVING_n_of_main_predicates	
0	431 (43%)
1	207 (21%)
2	183 (18%)
3	179 (18%)
ORDER_BY_n_of_columns	
0	275 (28%)
1	352 (35%)
2	248 (25%)
3	95 (9.5%)
4	15 (1.5%)
5	6 (0.6%)
6	5 (0.5%)
7	3 (0.3%)
8	1 (0.1%)
limit	[1, 504, 998] 501 / 291
offset	[0, 39, 998] 257 / 323