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

Scenario: 1-6jp_ujjp__g__h0ss_h0nss

Scale factor: 2000 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, 144] 15 / 21
SELECT_n_of_non_aggr_func___DAY	
0	878 (88%)
1	86 (8.6%)
2	23 (2.3%)
3	9 (0.9%)
4	2 (0.2%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___DOW	
0	834 (83%)
1	131 (13%)
2	22 (2.2%)
3	9 (0.9%)
4	1 (0.1%)

Query parameter or value	Descriptive statistics
5	3 (0.3%)
SELECT_n_of_non_aggr_func___LOG	
0	861 (86%)
1	104 (10%)
2	14 (1.4%)
3	14 (1.4%)
4	3 (0.3%)
5	2 (0.2%)
6	2 (0.2%)
SELECT_n_of_non_aggr_func___LOWER	
0	632 (63%)
1	250 (25%)
12	1 (0.1%)
2	72 (7.2%)
3	25 (2.5%)
4	12 (1.2%)
5	4 (0.4%)
6	2 (0.2%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func___LTRIM	
0	640 (64%)
1	241 (24%)
2	74 (7.4%)
3	22 (2.2%)
4	12 (1.2%)
5	7 (0.7%)
6	4 (0.4%)
SELECT_n_of_non_aggr_func___MONTH	
0	800 (80%)
1	129 (13%)
2	36 (3.6%)
3	21 (2.1%)
4	8 (0.8%)
5	2 (0.2%)
6	2 (0.2%)
9	2 (0.2%)
SELECT_n_of_non_aggr_func___ROUND	
0	921 (92%)
1	71 (7.1%)
2	8 (0.8%)
SELECT_n_of_non_aggr_func___RTRIM	
0	623 (62%)
1	254 (25%)
2	77 (7.7%)
3	22 (2.2%)
4	14 (1.4%)

Query parameter or value	Descriptive statistics
5	3 (0.3%)
6	6 (0.6%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___SQRT	
0	787 (79%)
1	150 (15%)
2	36 (3.6%)
3	13 (1.3%)
4	3 (0.3%)
5	7 (0.7%)
6	2 (0.2%)
7	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	630 (63%)
1	255 (26%)
2	69 (6.9%)
3	23 (2.3%)
4	9 (0.9%)
5	9 (0.9%)
6	4 (0.4%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___TRUNC	
0	927 (93%)
1	64 (6.4%)
2	7 (0.7%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___UPPER	
0	635 (64%)
1	242 (24%)
2	68 (6.8%)
3	28 (2.8%)
4	18 (1.8%)
5	1 (0.1%)
6	6 (0.6%)
7	1 (0.1%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func___YEAR	
0	815 (82%)
1	135 (14%)
2	28 (2.8%)
3	9 (0.9%)
4	4 (0.4%)
5	6 (0.6%)
6	2 (0.2%)
8	1 (0.1%)

Query parameter or value	Descriptive statistics
SELECT_n_of_non_aggr_func___FLOOR	
0	915 (92%)
1	73 (7.3%)
2	10 (1.0%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___ABS	
0	907 (91%)
1	80 (8.0%)
2	10 (1.0%)
3	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0, 3, 65] 5 / 7
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 12.00] 0.85 / 1.96
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 11.00] 0.91 / 1.94
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 12.00] 0.90 / 2.02
SELECT_n_of_aggr_func___AVG	
0	935 (94%)
1	47 (4.7%)
2	13 (1.3%)
3	4 (0.4%)
5	1 (0.1%)
SELECT_n_of_aggr_func___SUM	
0	929 (93%)
1	41 (4.1%)
2	22 (2.2%)
3	7 (0.7%)
4	1 (0.1%)
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 13.00] 0.79 / 1.87
SELECT_n_of_all_aggr_func	[0.0, 3.0, 14.0] 3.6 / 3.2
FROM_n_of_join_paths	
1	158 (16%)
2	170 (17%)
3	159 (16%)
4	183 (18%)
5	174 (17%)
6	156 (16%)
FROM_n_of_super_joins___FULL	
0	519 (52%)
1	326 (33%)
2	127 (13%)
3	27 (2.7%)
4	1 (0.1%)
FROM_n_of_super_joins___LEFT	
0	567 (57%)
1	300 (30%)

Query parameter or value	Descriptive statistics
2	112 (11%)
3	18 (1.8%)
4	3 (0.3%)
FROM_n_of_super_joins___RIGHT	
0	545 (55%)
1	323 (32%)
2	106 (11%)
3	25 (2.5%)
4	1 (0.1%)
FROM_n_of_joins___INNER	[0.00, 3.00, 11.00] 3.23 / 2.34
FROM_n_of_joins___RIGHT	[0.00, 3.00, 12.00] 3.33 / 2.50
FROM_n_of_processed_rows	[5, 18,530,027,853, 89,600,166,852] 23,646,260,943 / 19,655,740,233
WHERE_n_of_predicates	[0.0, 3.0, 12.0] 3.9 / 3.0
WHERE_n_of_attribs_of_type___character	
0	658 (66%)
1	249 (25%)
2	68 (6.8%)
3	20 (2.0%)
4	5 (0.5%)
WHERE_n_of_attribs_of_type___character_varying	
0	339 (34%)
1	296 (30%)
2	197 (20%)
3	107 (11%)
4	37 (3.7%)
5	16 (1.6%)
6	6 (0.6%)
7	2 (0.2%)
WHERE_n_of_attribs_of_type___date	
0	795 (80%)
1	169 (17%)
2	30 (3.0%)
3	6 (0.6%)
WHERE_n_of_attribs_of_type___numeric	
0	570 (57%)
1	268 (27%)
2	107 (11%)
3	43 (4.3%)
4	9 (0.9%)
5	3 (0.3%)
WHERE_n_of_attribs_of_type___integer	
0	363 (36%)
1	298 (30%)
2	178 (18%)
3	93 (9.3%)

Query parameter or value	Descriptive statistics
4	49 (4.9%)
5	15 (1.5%)
6	2 (0.2%)
7	2 (0.2%)
WHERE_n_of_pkey_attribs	
0	537 (54%)
1	290 (29%)
2	120 (12%)
3	44 (4.4%)
4	8 (0.8%)
5	1 (0.1%)
WHERE_n_of_connect_OR	[0.00, 1.00, 10.00] 2.22 / 2.40
WHERE_n_of_operators___greater_or_less	
0	323 (32%)
1	268 (27%)
2	217 (22%)
3	99 (9.9%)
4	55 (5.5%)
5	22 (2.2%)
6	14 (1.4%)
7	2 (0.2%)
WHERE_n_of_operators___in	
0	522 (52%)
1	305 (31%)
2	117 (12%)
3	46 (4.6%)
4	8 (0.8%)
6	2 (0.2%)
WHERE_n_of_operators___between	
0	520 (52%)
1	293 (29%)
2	117 (12%)
3	59 (5.9%)
4	8 (0.8%)
5	3 (0.3%)
WHERE_n_of_operators___like	
0	774 (77%)
1	188 (19%)
2	38 (3.8%)
WHERE_n_of_non_aggr_func___DAY	
0	966 (97%)
1	32 (3.2%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___TRUNC	
0	947 (95%)
1	53 (5.3%)

Query parameter or value	Descriptive statistics
WHERE_n_of_non_aggr_func___SQRT	
0	883 (88%)
1	111 (11%)
2	6 (0.6%)
WHERE_n_of_non_aggr_func___LOG	
0	912 (91%)
1	83 (8.3%)
2	5 (0.5%)
WHERE_n_of_non_aggr_func___DOW	
0	967 (97%)
1	33 (3.3%)
WHERE_n_of_non_aggr_func___ROUND	
0	948 (95%)
1	49 (4.9%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func___FLOOR	
0	963 (96%)
1	36 (3.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___ABS	
0	959 (96%)
1	39 (3.9%)
2	1 (0.1%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func___MONTH	
0	968 (97%)
1	30 (3.0%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___YEAR	
0	966 (97%)
1	34 (3.4%)
WHERE_n_of_all_non_aggr_func	
0	620 (62%)
1	257 (26%)
2	89 (8.9%)
3	31 (3.1%)
4	2 (0.2%)
7	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 12.00] 3.02 / 2.84
HAVING_n_of_main_predicates	
0	429 (43%)
1	193 (19%)
2	187 (19%)
3	191 (19%)
ORDER_BY_n_of_columns	
0	299 (30%)

Query parameter or value	Descriptive statistics
1	318 (32%)
2	256 (26%)
3	85 (8.5%)
4	21 (2.1%)
5	13 (1.3%)
6	5 (0.5%)
7	1 (0.1%)
8	2 (0.2%)
limit	[1, 491, 1,000] 498 / 287
offset	[0, 75, 999] 267 / 326