

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

Scenario: 1-6jp_ujpp__g__h0-3ss_h0-3nss

Scale factor: 200 GB

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

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, 194] 15 / 23
SELECT_n_of_non_aggr_func__LOWER	
0	613 (61%)
1	283 (28%)
10	1 (0.1%)
2	66 (6.6%)
3	16 (1.6%)
4	13 (1.3%)
5	5 (0.5%)
6	2 (0.2%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__SUBSTR	[0.00, 0.00, 10.00] 0.60 / 1.08
SELECT_n_of_non_aggr_func__ABS	
0	927 (93%)
1	63 (6.3%)

Query parameter or value	Descriptive statistics
2	9 (0.9%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__RTRIM	
0	649 (65%)
1	243 (24%)
12	1 (0.1%)
2	75 (7.5%)
3	15 (1.5%)
4	11 (1.1%)
5	3 (0.3%)
6	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__LTRIM	
0	595 (60%)
1	285 (29%)
2	66 (6.6%)
3	31 (3.1%)
4	10 (1.0%)
5	7 (0.7%)
6	3 (0.3%)
7	3 (0.3%)
SELECT_n_of_non_aggr_func__UPPER	
0	634 (63%)
1	239 (24%)
2	77 (7.7%)
3	30 (3.0%)
4	11 (1.1%)
5	5 (0.5%)
7	2 (0.2%)
8	2 (0.2%)
SELECT_n_of_non_aggr_func__TRUNC	
0	926 (93%)
1	59 (5.9%)
2	11 (1.1%)
3	4 (0.4%)
SELECT_n_of_non_aggr_func__LOG	
0	848 (85%)
1	120 (12%)
2	19 (1.9%)
3	7 (0.7%)
4	2 (0.2%)
5	2 (0.2%)
6	2 (0.2%)
SELECT_n_of_non_aggr_func__DAY	
0	905 (91%)
1	73 (7.3%)

Query parameter or value	Descriptive statistics
2	12 (1.2%)
3	6 (0.6%)
4	1 (0.1%)
5	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___DOW	
0	858 (86%)
1	104 (10%)
2	23 (2.3%)
3	8 (0.8%)
4	5 (0.5%)
5	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___YEAR	
0	828 (83%)
1	132 (13%)
2	22 (2.2%)
3	9 (0.9%)
4	7 (0.7%)
5	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___MONTH	[0.00, 0.00, 10.00] 0.28 / 0.83
SELECT_n_of_non_aggr_func___SQRT	
0	817 (82%)
1	128 (13%)
2	24 (2.4%)
3	13 (1.3%)
4	10 (1.0%)
5	6 (0.6%)
6	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___FLOOR	
0	926 (93%)
1	61 (6.1%)
2	12 (1.2%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___ROUND	
0	931 (93%)
1	60 (6.0%)
2	8 (0.8%)
3	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 72.0] 4.6 / 7.0
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 11.00] 0.76 / 1.75
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 12.00] 0.88 / 1.96
SELECT_n_of_aggr_func___SUM	
0	940 (94%)

Query parameter or value	Descriptive statistics
1	35 (3.5%)
2	20 (2.0%)
3	3 (0.3%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 12.00] 0.85 / 1.95
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 10.00] 0.89 / 1.96
SELECT_n_of_aggr_func__AVG	
0	930 (93%)
1	46 (4.6%)
2	16 (1.6%)
3	7 (0.7%)
4	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.0, 3.0, 13.0] 3.6 / 3.1
FROM_n_of_join_paths	
1	170 (17%)
2	179 (18%)
3	175 (18%)
4	156 (16%)
5	149 (15%)
6	171 (17%)
FROM_n_of_super_joins__LEFT	
0	552 (55%)
1	306 (31%)
2	106 (11%)
3	33 (3.3%)
4	3 (0.3%)
FROM_n_of_super_joins__RIGHT	
0	553 (55%)
1	323 (32%)
2	102 (10%)
3	20 (2.0%)
4	2 (0.2%)
FROM_n_of_super_joins__FULL	
0	543 (54%)
1	323 (32%)
2	111 (11%)
3	21 (2.1%)
4	2 (0.2%)
FROM_n_of_joins__INNER	[0.00, 3.00, 12.00] 3.23 / 2.45
FROM_n_of_joins__RIGHT	[0.00, 3.00, 15.00] 3.40 / 2.64
FROM_n_of_processed_rows	[5, 1,860,018,464, 8,810,110,684] 2,319,006,512 / 1,859,487,494
WHERE_n_of_predicates	[0.00, 3.00, 12.00] 3.80 / 2.92
WHERE_n_of_attris_of_type__character_varying	
0	368 (37%)

Query parameter or value	Descriptive statistics
1	293 (29%)
2	183 (18%)
3	94 (9.4%)
4	43 (4.3%)
5	15 (1.5%)
6	4 (0.4%)
WHERE_n_of_attribs_of_type__integer	
0	366 (37%)
1	290 (29%)
2	186 (19%)
3	90 (9.0%)
4	50 (5.0%)
5	13 (1.3%)
6	5 (0.5%)
WHERE_n_of_attribs_of_type__character	
0	662 (66%)
1	253 (25%)
2	60 (6.0%)
3	20 (2.0%)
4	4 (0.4%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type__numeric	
0	581 (58%)
1	273 (27%)
2	100 (10%)
3	32 (3.2%)
4	11 (1.1%)
5	3 (0.3%)
WHERE_n_of_attribs_of_type__date	
0	764 (76%)
1	197 (20%)
2	34 (3.4%)
3	5 (0.5%)
WHERE_n_of_pkey_attribs	
0	515 (52%)
1	304 (30%)
2	130 (13%)
3	36 (3.6%)
4	13 (1.3%)
5	2 (0.2%)
WHERE_n_of_connect_OR	[0.00, 1.00, 10.00] 2.18 / 2.40
WHERE_n_of_operators__between	
0	524 (52%)
1	313 (31%)
2	111 (11%)
3	41 (4.1%)

Query parameter or value	Descriptive statistics
4	11 (1.1%)
WHERE_n_of_operators__like	
0	788 (79%)
1	179 (18%)
2	30 (3.0%)
3	3 (0.3%)
WHERE_n_of_operators__greater_or_less	
0	347 (35%)
1	269 (27%)
2	169 (17%)
3	113 (11%)
4	59 (5.9%)
5	29 (2.9%)
6	14 (1.4%)
WHERE_n_of_operators__in	
0	522 (52%)
1	303 (30%)
2	126 (13%)
3	34 (3.4%)
4	13 (1.3%)
5	2 (0.2%)
WHERE_n_of_non_aggr_func__SQRT	
0	877 (88%)
1	109 (11%)
2	13 (1.3%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func__DOW	
0	956 (96%)
1	43 (4.3%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__LOG	
0	920 (92%)
1	75 (7.5%)
2	5 (0.5%)
WHERE_n_of_non_aggr_func__ROUND	
0	967 (97%)
1	31 (3.1%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__MONTH	
0	968 (97%)
1	32 (3.2%)
WHERE_n_of_non_aggr_func__TRUNC	
0	959 (96%)
1	41 (4.1%)
WHERE_n_of_non_aggr_func__FLOOR	
0	962 (96%)

Query parameter or value	Descriptive statistics
1	35 (3.5%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func__ABS	
0	960 (96%)
1	40 (4.0%)
WHERE_n_of_non_aggr_func__YEAR	
0	961 (96%)
1	35 (3.5%)
2	4 (0.4%)
WHERE_n_of_non_aggr_func__DAY	
0	958 (96%)
1	41 (4.1%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	625 (63%)
1	254 (25%)
2	87 (8.7%)
3	23 (2.3%)
4	9 (0.9%)
5	2 (0.2%)
GROUP_BY_n_of_columns	[0.00, 3.00, 12.00] 3.10 / 2.79
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.51 / 2.59
HAVING_n_of_main_predicates__non_scalar_subquery	
0	425 (43%)
1	185 (19%)
2	196 (20%)
3	194 (19%)
HAVING_n_of_main_predicates__scalar_subquery	
0	397 (40%)
1	200 (20%)
2	185 (19%)
3	218 (22%)
HAVING_n_of_subqueries__non_scalar_subquery	
0	425 (43%)
1	185 (19%)
2	196 (20%)
3	194 (19%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 21.0] 2.6 / 3.5
HAVING_n_of_processed_rows_by_subquery	[0, 3, 120, 018, 464, 132, 151, 660, 260] 8,888,596,216 / 14,601,201,710
ORDER_BY_n_of_columns	
0	277 (28%)
1	344 (34%)
2	257 (26%)
3	93 (9.3%)
4	15 (1.5%)

Query parameter or value	Descriptive statistics
5	10 (1.0%)
6	4 (0.4%)
limit	[1, 521, 1,000] 509 / 291
offset	[0, 0, 1,000] 237 / 316