

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

Scenario: 1-6jp_rjpp__g__h0ss_h0nss

Scale factor: 2000 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, 170] 16 / 26
SELECT_n_of_non_aggr_func__LTRIM	
0	616 (62%)
1	261 (26%)
10	1 (0.1%)
2	65 (6.5%)
3	35 (3.5%)
4	10 (1.0%)
5	4 (0.4%)
6	7 (0.7%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__RTRIM	
0	616 (62%)
1	245 (25%)
2	90 (9.0%)

Query parameter or value	Descriptive statistics
3	24 (2.4%)
4	12 (1.2%)
5	4 (0.4%)
6	6 (0.6%)
7	2 (0.2%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func__LOG	
0	869 (87%)
1	101 (10%)
2	16 (1.6%)
3	7 (0.7%)
4	4 (0.4%)
5	1 (0.1%)
6	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__YEAR	
0	833 (83%)
1	109 (11%)
2	28 (2.8%)
3	19 (1.9%)
4	5 (0.5%)
5	3 (0.3%)
6	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__SUBSTR	[0.00, 0.00, 9.00] 0.57 / 1.03
SELECT_n_of_non_aggr_func__LOWER	
0	608 (61%)
1	262 (26%)
2	88 (8.8%)
3	23 (2.3%)
4	5 (0.5%)
5	5 (0.5%)
6	5 (0.5%)
7	3 (0.3%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func__UPPER	
0	597 (60%)
1	281 (28%)
2	73 (7.3%)
3	21 (2.1%)
4	16 (1.6%)
5	6 (0.6%)
6	3 (0.3%)
7	2 (0.2%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__DAY	

Query parameter or value	Descriptive statistics
0	879 (88%)
1	95 (9.5%)
2	13 (1.3%)
3	9 (0.9%)
4	1 (0.1%)
5	2 (0.2%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__MONTH	[0.00, 0.00, 10.00] 0.34 / 0.96
SELECT_n_of_non_aggr_func__ABS	
0	919 (92%)
1	71 (7.1%)
2	7 (0.7%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func__DOW	
0	834 (83%)
1	125 (13%)
2	21 (2.1%)
3	12 (1.2%)
4	7 (0.7%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__FLOOR	
0	926 (93%)
1	58 (5.8%)
2	10 (1.0%)
3	5 (0.5%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__SQRT	
0	803 (80%)
1	138 (14%)
2	32 (3.2%)
3	12 (1.2%)
4	7 (0.7%)
5	5 (0.5%)
6	3 (0.3%)
SELECT_n_of_non_aggr_func__TRUNC	
0	935 (94%)
1	54 (5.4%)
2	7 (0.7%)
3	2 (0.2%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__ROUND	
0	928 (93%)
1	58 (5.8%)
2	10 (1.0%)
3	2 (0.2%)

Query parameter or value	Descriptive statistics
4	2 (0.2%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 66.0] 4.9 / 7.7
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 12.00] 0.77 / 1.79
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 13.00] 0.89 / 2.00
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 14.00] 0.87 / 2.00
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 13.00] 0.78 / 1.91
SELECT_n_of_aggr_func__AVG	
0	949 (95%)
1	30 (3.0%)
2	15 (1.5%)
3	4 (0.4%)
4	2 (0.2%)
SELECT_n_of_aggr_func__SUM	
0	935 (94%)
1	44 (4.4%)
2	16 (1.6%)
3	4 (0.4%)
4	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.0, 3.0, 16.0] 3.5 / 3.2
FROM_n_of_join_paths	
1	171 (17%)
2	178 (18%)
3	160 (16%)
4	151 (15%)
5	161 (16%)
6	179 (18%)
FROM_n_of_super_joins__FULL	
0	548 (55%)
1	301 (30%)
2	120 (12%)
3	25 (2.5%)
4	6 (0.6%)
FROM_n_of_super_joins__RIGHT	
0	538 (54%)
1	327 (33%)
2	107 (11%)
3	22 (2.2%)
4	6 (0.6%)
FROM_n_of_super_joins__LEFT	
0	573 (57%)
1	280 (28%)
2	122 (12%)
3	23 (2.3%)
4	1 (0.1%)
5	1 (0.1%)
FROM_n_of_joins__INNER	[0.00, 3.00, 13.00] 3.34 / 2.48

Query parameter or value	Descriptive statistics
FROM_n_of_joins__RIGHT	[0.00, 3.00, 13.00] 3.21 / 2.47
FROM_n_of_processed_rows	[5, 18,880,027,852, 87,820,166,894] 24,326,182,414 / 20,232,405,824
WHERE_n_of_predicates	[0.0, 4.0, 13.0] 3.9 / 3.0
WHERE_n_of_attribs_of_type__character_varying	
0	364 (36%)
1	277 (28%)
2	184 (18%)
3	108 (11%)
4	46 (4.6%)
5	18 (1.8%)
6	3 (0.3%)
WHERE_n_of_attribs_of_type__numeric	
0	556 (56%)
1	268 (27%)
2	119 (12%)
3	42 (4.2%)
4	10 (1.0%)
5	5 (0.5%)
WHERE_n_of_attribs_of_type__date	
0	768 (77%)
1	187 (19%)
2	36 (3.6%)
3	6 (0.6%)
4	3 (0.3%)
WHERE_n_of_attribs_of_type__integer	
0	365 (37%)
1	279 (28%)
2	193 (19%)
3	114 (11%)
4	30 (3.0%)
5	12 (1.2%)
6	5 (0.5%)
7	2 (0.2%)
WHERE_n_of_attribs_of_type__character	
0	643 (64%)
1	269 (27%)
2	64 (6.4%)
3	22 (2.2%)
4	2 (0.2%)
WHERE_n_of_pkey_attribs	
0	523 (52%)
1	302 (30%)
2	123 (12%)
3	46 (4.6%)
4	4 (0.4%)

Query parameter or value	Descriptive statistics
5	2 (0.2%)
WHERE_n_of_connect_OR	[0.00, 2.00, 11.00] 2.36 / 2.53
WHERE_n_of_operators___greater_or_less	
0	335 (34%)
1	258 (26%)
2	187 (19%)
3	114 (11%)
4	65 (6.5%)
5	24 (2.4%)
6	13 (1.3%)
7	3 (0.3%)
8	1 (0.1%)
WHERE_n_of_operators___in	
0	527 (53%)
1	293 (29%)
2	119 (12%)
3	39 (3.9%)
4	16 (1.6%)
5	5 (0.5%)
7	1 (0.1%)
WHERE_n_of_operators___between	
0	501 (50%)
1	321 (32%)
2	135 (14%)
3	31 (3.1%)
4	12 (1.2%)
WHERE_n_of_operators___like	
0	792 (79%)
1	178 (18%)
2	27 (2.7%)
3	3 (0.3%)
WHERE_n_of_non_aggr_func___ROUND	
0	950 (95%)
1	50 (5.0%)
WHERE_n_of_non_aggr_func___ABS	
0	948 (95%)
1	51 (5.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___FLOOR	
0	948 (95%)
1	50 (5.0%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___SQRT	
0	875 (88%)
1	115 (12%)
2	9 (0.9%)

Query parameter or value	Descriptive statistics
3	1 (0.1%)
WHERE_n_of_non_aggr_func__YEAR	
0	954 (95%)
1	45 (4.5%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__LOG	
0	928 (93%)
1	66 (6.6%)
2	5 (0.5%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func__DOW	
0	959 (96%)
1	40 (4.0%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DAY	
0	960 (96%)
1	40 (4.0%)
WHERE_n_of_non_aggr_func__MONTH	
0	952 (95%)
1	45 (4.5%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func__TRUNC	
0	960 (96%)
1	39 (3.9%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	591 (59%)
1	269 (27%)
2	102 (10%)
3	32 (3.2%)
4	6 (0.6%)
GROUP_BY_n_of_columns	[0.00, 2.50, 12.00] 3.11 / 2.87
HAVING_n_of_main_predicates	
0	431 (43%)
1	176 (18%)
2	205 (21%)
3	188 (19%)
ORDER_BY_n_of_columns	
0	283 (28%)
1	331 (33%)
2	251 (25%)
3	92 (9.2%)
4	28 (2.8%)
5	10 (1.0%)
6	2 (0.2%)
7	2 (0.2%)

Query parameter or value	Descriptive statistics
8	1 (0.1%)
limit	[1, 522, 1,000] 514 / 290
offset	[0, 0, 998] 250 / 325