

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

Scenario: 1-6jp_rjpp__g__h0-3ss_h0nss

Scale factor: 0.01 GB

The 1000-query pack was generated on: 2023-11-07

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, 185] 14 / 22
SELECT_n_of_non_aggr_func__FLOOR	
0	910 (91%)
1	75 (7.5%)
2	9 (0.9%)
3	5 (0.5%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__LOWER	
0	621 (62%)
1	267 (27%)
2	79 (7.9%)
3	16 (1.6%)
4	11 (1.1%)
5	2 (0.2%)
6	1 (0.1%)

Query parameter or value	Descriptive statistics
7	3 (0.3%)
SELECT__n_of__non__aggr__func__RTRIM	[0.00, 0.00, 11.00] 0.58 / 1.00
SELECT__n_of__non__aggr__func__SQRT	
0	832 (83%)
1	125 (13%)
2	25 (2.5%)
3	9 (0.9%)
4	2 (0.2%)
5	2 (0.2%)
6	3 (0.3%)
7	2 (0.2%)
SELECT__n_of__non__aggr__func__SUBSTR	[0.00, 0.00, 9.00] 0.54 / 0.94
SELECT__n_of__non__aggr__func__ABS	
0	921 (92%)
1	71 (7.1%)
2	5 (0.5%)
3	3 (0.3%)
SELECT__n_of__non__aggr__func__LTRIM	
0	644 (64%)
1	253 (25%)
2	67 (6.7%)
3	20 (2.0%)
4	9 (0.9%)
5	2 (0.2%)
6	3 (0.3%)
7	2 (0.2%)
SELECT__n_of__non__aggr__func__UPPER	[0.00, 0.00, 9.00] 0.59 / 1.02
SELECT__n_of__non__aggr__func__MONTH	[0.00, 0.00, 12.00] 0.31 / 0.93
SELECT__n_of__non__aggr__func__YEAR	
0	844 (84%)
1	117 (12%)
2	21 (2.1%)
3	8 (0.8%)
4	4 (0.4%)
5	2 (0.2%)
6	3 (0.3%)
7	1 (0.1%)
SELECT__n_of__non__aggr__func__DAY	
0	886 (89%)
1	91 (9.1%)
2	16 (1.6%)
3	3 (0.3%)
4	2 (0.2%)
5	1 (0.1%)
6	1 (0.1%)
SELECT__n_of__non__aggr__func__LOG	

Query parameter or value	Descriptive statistics
0	866 (87%)
1	112 (11%)
2	14 (1.4%)
3	6 (0.6%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___DOW	
0	871 (87%)
1	98 (9.8%)
2	16 (1.6%)
3	7 (0.7%)
4	6 (0.6%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___ROUND	
0	935 (94%)
1	56 (5.6%)
2	5 (0.5%)
3	2 (0.2%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___TRUNC	
0	919 (92%)
1	62 (6.2%)
2	17 (1.7%)
3	2 (0.2%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 67.0] 4.5 / 7.0
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 12.00] 0.81 / 1.83
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 11.00] 0.79 / 1.85
SELECT_n_of_aggr_func___SUM	
0	927 (93%)
1	43 (4.3%)
2	19 (1.9%)
3	7 (0.7%)
4	3 (0.3%)
5	1 (0.1%)
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 13.00] 0.87 / 2.00
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 15.00] 0.84 / 1.90
SELECT_n_of_aggr_func___AVG	
0	940 (94%)
1	39 (3.9%)
2	16 (1.6%)
3	4 (0.4%)
4	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.0, 3.0, 15.0] 3.5 / 3.1
FROM_n_of_join_paths	
1	162 (16%)
2	173 (17%)
3	186 (19%)

Query parameter or value	Descriptive statistics
4	177 (18%)
5	152 (15%)
6	150 (15%)
FROM_n_of_super_joins__LEFT	
0	554 (55%)
1	302 (30%)
2	121 (12%)
3	20 (2.0%)
4	3 (0.3%)
FROM_n_of_super_joins__RIGHT	
0	553 (55%)
1	322 (32%)
2	103 (10%)
3	20 (2.0%)
4	2 (0.2%)
FROM_n_of_super_joins__FULL	
0	551 (55%)
1	317 (32%)
2	108 (11%)
3	22 (2.2%)
4	2 (0.2%)
FROM_n_of_joins__INNER	[0.00, 3.00, 13.00] 3.19 / 2.47
FROM_n_of_joins__RIGHT	[0.00, 3.00, 12.00] 3.24 / 2.42
FROM_n_of_processed_rows	[5, 93,203, 434,985] 117,971 / 97,033
WHERE_n_of_predicates	[0.00, 3.00, 12.00] 3.62 / 2.94
WHERE_n_of_attribs_of_type__character	
0	676 (68%)
1	237 (24%)
2	73 (7.3%)
3	14 (1.4%)
WHERE_n_of_attribs_of_type__character_varying	
0	389 (39%)
1	305 (31%)
2	165 (17%)
3	80 (8.0%)
4	36 (3.6%)
5	19 (1.9%)
6	5 (0.5%)
7	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	782 (78%)
1	182 (18%)
2	28 (2.8%)
3	6 (0.6%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type__integer	

Query parameter or value	Descriptive statistics
0	395 (40%)
1	280 (28%)
2	181 (18%)
3	94 (9.4%)
4	36 (3.6%)
5	10 (1.0%)
6	4 (0.4%)
WHERE_n_of_attribs_of_type__numeric	
0	586 (59%)
1	271 (27%)
2	90 (9.0%)
3	36 (3.6%)
4	12 (1.2%)
5	3 (0.3%)
6	2 (0.2%)
WHERE_n_of_pkey_attribs	
0	553 (55%)
1	289 (29%)
2	122 (12%)
3	30 (3.0%)
4	4 (0.4%)
5	2 (0.2%)
WHERE_n_of_connect_OR	[0.00, 1.00, 11.00] 2.04 / 2.42
WHERE_n_of_operators__between	
0	545 (55%)
1	303 (30%)
2	100 (10%)
3	39 (3.9%)
4	9 (0.9%)
5	4 (0.4%)
WHERE_n_of_operators__greater_or_less	
0	347 (35%)
1	282 (28%)
2	175 (18%)
3	96 (9.6%)
4	64 (6.4%)
5	24 (2.4%)
6	9 (0.9%)
7	3 (0.3%)
WHERE_n_of_operators__in	
0	564 (56%)
1	290 (29%)
2	90 (9.0%)
3	38 (3.8%)
4	15 (1.5%)
5	2 (0.2%)

Query parameter or value	Descriptive statistics
7	1 (0.1%)
WHERE_n_of_operators_like	
0	797 (80%)
1	176 (18%)
2	25 (2.5%)
3	2 (0.2%)
WHERE_n_of_non_aggr_func_ABS	
0	958 (96%)
1	41 (4.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_FLOOR	
0	962 (96%)
1	37 (3.7%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_MONTH	
0	969 (97%)
1	31 (3.1%)
WHERE_n_of_non_aggr_func_SQRT	
0	879 (88%)
1	103 (10%)
2	16 (1.6%)
3	2 (0.2%)
WHERE_n_of_non_aggr_func_LOG	
0	926 (93%)
1	71 (7.1%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func_TRUNC	
0	954 (95%)
1	46 (4.6%)
WHERE_n_of_non_aggr_func_DAY	
0	967 (97%)
1	29 (2.9%)
2	4 (0.4%)
WHERE_n_of_non_aggr_func_DOW	
0	955 (96%)
1	43 (4.3%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func_YEAR	
0	949 (95%)
1	50 (5.0%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_ROUND	
0	960 (96%)
1	39 (3.9%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	

Query parameter or value	Descriptive statistics
0	617 (62%)
1	267 (27%)
2	77 (7.7%)
3	27 (2.7%)
4	9 (0.9%)
5	2 (0.2%)
6	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 3.00, 13.00] 3.08 / 2.81
HAVING_n_of_main_predicates	
0	291 (29%)
1	121 (12%)
2	126 (13%)
3	193 (19%)
4	122 (12%)
5	88 (8.8%)
6	59 (5.9%)
HAVING_n_of_main_predicates__scalar_subquery	
0	452 (45%)
1	197 (20%)
2	164 (16%)
3	187 (19%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 27.0] 2.3 / 3.7
HAVING_n_of_processed_rows_by_subqueries	[0, 8,283, 5,187,870] 279,111 / 604,688
ORDER_BY_n_of_columns	
0	302 (30%)
1	345 (35%)
2	226 (23%)
3	91 (9.1%)
4	19 (1.9%)
5	7 (0.7%)
6	8 (0.8%)
7	2 (0.2%)
limit	[1, 472, 999] 482 / 291
offset	[0, 23, 1,000] 243 / 318