

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

## Scenario: 1-6jp\_ujpp\_\_g\_\_h0ss\_h0nss

## Scale factor: 50 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, 9, 158] 16 / 23
SELECT_n_of_non_aggr_func__LOWER	[0.00, 0.00, 13.00] 0.59 / 1.06
SELECT_n_of_non_aggr_func__LTRIM	
0	621 (62%)
1	243 (24%)
2	96 (9.6%)
3	21 (2.1%)
4	8 (0.8%)
5	7 (0.7%)
6	3 (0.3%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__DAY	
0	882 (88%)
1	96 (9.6%)
2	13 (1.3%)

Query parameter or value	Descriptive statistics
3	6 (0.6%)
4	3 (0.3%)
SELECT__n_of__non__aggr__func__FLOOR	
0	923 (92%)
1	65 (6.5%)
2	11 (1.1%)
3	1 (0.1%)
SELECT__n_of__non__aggr__func__MONTH	
0	802 (80%)
1	142 (14%)
2	31 (3.1%)
3	13 (1.3%)
4	3 (0.3%)
5	5 (0.5%)
6	3 (0.3%)
7	1 (0.1%)
SELECT__n_of__non__aggr__func__RTRIM	
0	640 (64%)
1	247 (25%)
2	73 (7.3%)
3	21 (2.1%)
4	9 (0.9%)
5	8 (0.8%)
6	2 (0.2%)
SELECT__n_of__non__aggr__func__SQRT	
0	817 (82%)
1	126 (13%)
2	33 (3.3%)
3	10 (1.0%)
4	8 (0.8%)
5	3 (0.3%)
6	3 (0.3%)
SELECT__n_of__non__aggr__func__UPPER	
0	624 (62%)
1	268 (27%)
2	67 (6.7%)
3	23 (2.3%)
4	8 (0.8%)
5	8 (0.8%)
6	1 (0.1%)
7	1 (0.1%)
SELECT__n_of__non__aggr__func__DOW	
0	857 (86%)
1	102 (10%)
2	26 (2.6%)
3	8 (0.8%)

Query parameter or value	Descriptive statistics
4	6 (0.6%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__YEAR	
0	827 (83%)
1	135 (14%)
2	21 (2.1%)
3	10 (1.0%)
4	4 (0.4%)
5	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	627 (63%)
1	267 (27%)
2	67 (6.7%)
3	17 (1.7%)
4	7 (0.7%)
5	9 (0.9%)
6	3 (0.3%)
7	1 (0.1%)
9	2 (0.2%)
SELECT_n_of_non_aggr_func__ABS	
0	923 (92%)
1	67 (6.7%)
2	9 (0.9%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__LOG	
0	863 (86%)
1	105 (11%)
2	23 (2.3%)
3	7 (0.7%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__ROUND	
0	909 (91%)
1	78 (7.8%)
2	12 (1.2%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__TRUNC	
0	939 (94%)
1	55 (5.5%)
2	5 (0.5%)
3	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 70.0] 4.6 / 6.8
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 10.00] 0.77 / 1.81
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 14.00] 0.86 / 2.01
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 13.00] 0.88 / 1.98

Query parameter or value	Descriptive statistics
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 14.00] 0.94 / 2.11
SELECT_n_of_aggr_func__SUM	
0	936 (94%)
1	38 (3.8%)
2	19 (1.9%)
3	6 (0.6%)
4	1 (0.1%)
SELECT_n_of_aggr_func__AVG	
0	920 (92%)
1	48 (4.8%)
2	19 (1.9%)
3	10 (1.0%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.0, 3.0, 15.0] 3.7 / 3.2
FROM_n_of_join_paths	
1	167 (17%)
2	159 (16%)
3	166 (17%)
4	171 (17%)
5	181 (18%)
6	156 (16%)
FROM_n_of_super_joins__RIGHT	
0	521 (52%)
1	329 (33%)
2	124 (12%)
3	25 (2.5%)
4	1 (0.1%)
FROM_n_of_super_joins__FULL	
0	561 (56%)
1	301 (30%)
2	117 (12%)
3	18 (1.8%)
4	3 (0.3%)
FROM_n_of_super_joins__LEFT	
0	530 (53%)
1	336 (34%)
2	110 (11%)
3	22 (2.2%)
4	2 (0.2%)
FROM_n_of_joins__INNER	[0.00, 3.00, 13.00] 3.22 / 2.42
FROM_n_of_joins__RIGHT	[0.00, 3.00, 13.00] 3.33 / 2.57
FROM_n_of_processed_rows	[5, 471,712,935, 2,165,059,023] 596,450,918 / 493,730,670
WHERE_n_of_predicates	[0.00, 3.00, 13.00] 3.85 / 2.97
WHERE_n_of_attrbs_of_type__character_varying	

Query parameter or value	Descriptive statistics
0	335 (34%)
1	294 (29%)
2	195 (20%)
3	102 (10%)
4	48 (4.8%)
5	17 (1.7%)
6	7 (0.7%)
7	1 (0.1%)
8	1 (0.1%)
WHERE_n_of_attribs_of_type__integer	
0	394 (39%)
1	285 (29%)
2	176 (18%)
3	96 (9.6%)
4	32 (3.2%)
5	15 (1.5%)
6	2 (0.2%)
WHERE_n_of_attribs_of_type__numeric	
0	592 (59%)
1	270 (27%)
2	95 (9.5%)
3	32 (3.2%)
4	10 (1.0%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	770 (77%)
1	183 (18%)
2	37 (3.7%)
3	9 (0.9%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type__character	
0	641 (64%)
1	244 (24%)
2	93 (9.3%)
3	18 (1.8%)
4	4 (0.4%)
WHERE_n_of_pkey_attribs	
0	560 (56%)
1	283 (28%)
2	124 (12%)
3	26 (2.6%)
4	7 (0.7%)
WHERE_n_of_connect_OR	[0.00, 1.00, 10.00] 2.18 / 2.42
WHERE_n_of_operators__greater_or_less	
0	357 (36%)
1	247 (25%)

Query parameter or value	Descriptive statistics
2	178 (18%)
3	118 (12%)
4	67 (6.7%)
5	23 (2.3%)
6	6 (0.6%)
7	3 (0.3%)
8	1 (0.1%)
WHERE_n_of_operators_in	
0	519 (52%)
1	294 (29%)
2	127 (13%)
3	47 (4.7%)
4	8 (0.8%)
5	5 (0.5%)
WHERE_n_of_operators_like	
0	779 (78%)
1	183 (18%)
2	35 (3.5%)
3	2 (0.2%)
4	1 (0.1%)
WHERE_n_of_operators_between	
0	519 (52%)
1	290 (29%)
2	133 (13%)
3	41 (4.1%)
4	15 (1.5%)
5	2 (0.2%)
WHERE_n_of_non_aggr_func_LOG	
0	913 (91%)
1	84 (8.4%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func_MONTH	
0	960 (96%)
1	39 (3.9%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_DAY	
0	964 (96%)
1	35 (3.5%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func_TRUNC	
0	960 (96%)
1	40 (4.0%)
WHERE_n_of_non_aggr_func_SQRT	
0	891 (89%)
1	99 (9.9%)
2	10 (1.0%)

Query parameter or value	Descriptive statistics
WHERE_n_of_non_aggr_func__FLOOR	
0	953 (95%)
1	45 (4.5%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__ROUND	
0	960 (96%)
1	40 (4.0%)
WHERE_n_of_non_aggr_func__YEAR	
0	956 (96%)
1	44 (4.4%)
WHERE_n_of_non_aggr_func__ABS	
0	962 (96%)
1	35 (3.5%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func__DOW	
0	954 (95%)
1	43 (4.3%)
2	3 (0.3%)
WHERE_n_of_all_non_aggr_func	
0	607 (61%)
1	274 (27%)
2	87 (8.7%)
3	28 (2.8%)
4	2 (0.2%)
5	1 (0.1%)
6	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 3.00, 12.00] 3.11 / 2.87
HAVING_n_of_main_predicates	
0	437 (44%)
1	194 (19%)
2	193 (19%)
3	176 (18%)
ORDER_BY_n_of_columns	
0	313 (31%)
1	292 (29%)
2	247 (25%)
3	106 (11%)
4	25 (2.5%)
5	7 (0.7%)
6	6 (0.6%)
7	4 (0.4%)
limit	[1, 503, 999] 497 / 293
offset	[0, 0, 998] 248 / 327