

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

## Scenario: 1-6jp\_rjpp\_\_g\_\_h0-3ss\_h0-3nss

## Scale factor: 50 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, 200] 16 / 25
SELECT_n_of_non_aggr_func__SUBSTR	
0	636 (64%)
1	254 (25%)
2	59 (5.9%)
3	24 (2.4%)
4	10 (1.0%)
5	9 (0.9%)
6	4 (0.4%)
7	1 (0.1%)
8	3 (0.3%)
SELECT_n_of_non_aggr_func__LTRIM	[0.00, 0.00, 11.00] 0.61 / 1.10
SELECT_n_of_non_aggr_func__DOW	
0	852 (85%)
1	110 (11%)

Query parameter or value	Descriptive statistics
2	24 (2.4%)
3	8 (0.8%)
4	3 (0.3%)
5	2 (0.2%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___UPPER	
0	615 (62%)
1	265 (27%)
2	81 (8.1%)
3	17 (1.7%)
4	10 (1.0%)
5	7 (0.7%)
6	2 (0.2%)
8	3 (0.3%)
SELECT_n_of_non_aggr_func___YEAR	
0	838 (84%)
1	113 (11%)
10	1 (0.1%)
2	21 (2.1%)
3	12 (1.2%)
4	10 (1.0%)
5	3 (0.3%)
6	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___RTRIM	
0	648 (65%)
1	245 (25%)
2	58 (5.8%)
3	27 (2.7%)
4	11 (1.1%)
5	7 (0.7%)
6	1 (0.1%)
7	2 (0.2%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___SQRT	
0	800 (80%)
1	145 (15%)
2	27 (2.7%)
3	11 (1.1%)
4	5 (0.5%)
5	7 (0.7%)
6	2 (0.2%)
7	3 (0.3%)
SELECT_n_of_non_aggr_func___LOWER	
0	626 (63%)
1	250 (25%)

Query parameter or value	Descriptive statistics
2	80 (8.0%)
3	18 (1.8%)
4	17 (1.7%)
5	6 (0.6%)
6	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___ABS	
0	912 (91%)
1	77 (7.7%)
2	6 (0.6%)
3	5 (0.5%)
SELECT_n_of_non_aggr_func___DAY	
0	893 (89%)
1	72 (7.2%)
2	24 (2.4%)
3	6 (0.6%)
4	2 (0.2%)
5	3 (0.3%)
SELECT_n_of_non_aggr_func___MONTH	
0	818 (82%)
1	120 (12%)
10	2 (0.2%)
2	32 (3.2%)
3	12 (1.2%)
4	9 (0.9%)
5	4 (0.4%)
6	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___ROUND	
0	913 (91%)
1	73 (7.3%)
2	10 (1.0%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___FLOOR	
0	918 (92%)
1	73 (7.3%)
2	8 (0.8%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___TRUNC	
0	933 (93%)
1	56 (5.6%)
2	9 (0.9%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___LOG	
0	861 (86%)

Query parameter or value	Descriptive statistics
1	108 (11%)
2	22 (2.2%)
3	5 (0.5%)
5	2 (0.2%)
6	1 (0.1%)
9	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 66.0] 4.8 / 7.6
SELECT_n_of_aggr_func__AVG	
0	940 (94%)
1	38 (3.8%)
2	15 (1.5%)
3	6 (0.6%)
6	1 (0.1%)
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 12.00] 0.84 / 1.90
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 13.00] 0.84 / 1.95
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 12.00] 0.83 / 1.93
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 10.00] 0.94 / 2.05
SELECT_n_of_aggr_func__SUM	
0	932 (93%)
1	41 (4.1%)
2	22 (2.2%)
3	3 (0.3%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.0, 3.0, 14.0] 3.6 / 3.1
FROM_n_of_join_paths	
1	170 (17%)
2	168 (17%)
3	166 (17%)
4	166 (17%)
5	164 (16%)
6	166 (17%)
FROM_n_of_super_joins__FULL	
0	552 (55%)
1	316 (32%)
2	103 (10%)
3	27 (2.7%)
4	2 (0.2%)
FROM_n_of_super_joins__LEFT	
0	524 (52%)
1	319 (32%)
2	133 (13%)
3	24 (2.4%)
FROM_n_of_super_joins__RIGHT	
0	553 (55%)
1	313 (31%)

Query parameter or value	Descriptive statistics
2	109 (11%)
3	19 (1.9%)
4	5 (0.5%)
5	1 (0.1%)
FROM_n_of_joins__INNER	[0.00, 3.00, 13.00] 3.33 / 2.52
FROM_n_of_joins__RIGHT	[0.00, 3.00, 13.00] 3.31 / 2.46
FROM_n_of_processed_rows	[5, 465,031,225, 2,295,098,497] 600,405,767 / 493,806,965
WHERE_n_of_predicates	[0.0, 3.0, 14.0] 3.8 / 3.0
WHERE_n_of_attribs_of_type__character_varying	
0	346 (35%)
1	297 (30%)
2	202 (20%)
3	100 (10%)
4	37 (3.7%)
5	15 (1.5%)
7	3 (0.3%)
WHERE_n_of_attribs_of_type__date	
0	794 (79%)
1	159 (16%)
2	38 (3.8%)
3	8 (0.8%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type__integer	
0	367 (37%)
1	293 (29%)
2	188 (19%)
3	95 (9.5%)
4	43 (4.3%)
5	11 (1.1%)
6	2 (0.2%)
7	1 (0.1%)
WHERE_n_of_attribs_of_type__numeric	
0	575 (58%)
1	278 (28%)
2	100 (10%)
3	32 (3.2%)
4	10 (1.0%)
5	5 (0.5%)
WHERE_n_of_attribs_of_type__character	
0	640 (64%)
1	269 (27%)
2	78 (7.8%)
3	11 (1.1%)
4	2 (0.2%)
WHERE_n_of_pkey_attribs	

Query parameter or value	Descriptive statistics
0	526 (53%)
1	320 (32%)
2	113 (11%)
3	35 (3.5%)
4	6 (0.6%)
WHERE_n_of_connect_OR	[0.00, 1.00, 12.00] 2.22 / 2.51
WHERE_n_of_operators___between	
0	542 (54%)
1	306 (31%)
2	100 (10%)
3	35 (3.5%)
4	15 (1.5%)
5	1 (0.1%)
6	1 (0.1%)
WHERE_n_of_operators___greater_or_less	
0	348 (35%)
1	240 (24%)
2	178 (18%)
3	121 (12%)
4	67 (6.7%)
5	27 (2.7%)
6	14 (1.4%)
7	4 (0.4%)
8	1 (0.1%)
WHERE_n_of_operators___in	
0	563 (56%)
1	267 (27%)
2	113 (11%)
3	46 (4.6%)
4	10 (1.0%)
5	1 (0.1%)
WHERE_n_of_operators___like	
0	776 (78%)
1	193 (19%)
2	27 (2.7%)
3	3 (0.3%)
4	1 (0.1%)
WHERE_n_of_non_aggr_func___LOG	
0	931 (93%)
1	63 (6.3%)
2	6 (0.6%)
WHERE_n_of_non_aggr_func___MONTH	
0	960 (96%)
1	35 (3.5%)
2	5 (0.5%)
WHERE_n_of_non_aggr_func___SQRT	

Query parameter or value	Descriptive statistics
0	884 (88%)
1	108 (11%)
2	8 (0.8%)
WHERE_n_of_non_aggr_func__YEAR	
0	972 (97%)
1	27 (2.7%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__ROUND	
0	966 (97%)
1	34 (3.4%)
WHERE_n_of_non_aggr_func__DAY	
0	958 (96%)
1	41 (4.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__ABS	
0	956 (96%)
1	41 (4.1%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func__DOW	
0	958 (96%)
1	39 (3.9%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func__FLOOR	
0	969 (97%)
1	30 (3.0%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__TRUNC	
0	955 (96%)
1	44 (4.4%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	639 (64%)
1	244 (24%)
2	85 (8.5%)
3	25 (2.5%)
4	4 (0.4%)
5	3 (0.3%)
GROUP_BY_n_of_columns	[0.00, 2.00, 12.00] 2.93 / 2.76
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.36 / 2.59
HAVING_n_of_main_predicates__non_scalar_subquery	
0	453 (45%)
1	180 (18%)
2	183 (18%)
3	184 (18%)
HAVING_n_of_main_predicates__scalar_subquery	
0	427 (43%)

Query parameter or value	Descriptive statistics
<i>1</i>	204 (20%)
<i>2</i>	183 (18%)
<i>3</i>	186 (19%)
HAVING_n_of_subqueries__non_scalar_subquery	
<i>0</i>	453 (45%)
<i>1</i>	180 (18%)
<i>2</i>	183 (18%)
<i>3</i>	184 (18%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 21.0] 2.4 / 3.5
HAVING_n_of_processed_rows_by_subqueries	[0, 680,483,908, 20,671,045,447] 1,964,182,584 / 3,166,252,944
ORDER_BY_n_of_columns	[0.00, 1.00, 10.00] 1.28 / 1.19
limit	[2, 502, 997] 500 / 289
offset	[0, 22, 1,000] 247 / 319