

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

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

## Scale factor: 5 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, 9, 158] 16 / 22
SELECT_n_of_non_aggr_func__YEAR	
0	810 (81%)
1	135 (14%)
2	29 (2.9%)
3	12 (1.2%)
4	5 (0.5%)
5	4 (0.4%)
6	3 (0.3%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func__SQRT	
0	795 (80%)
1	153 (15%)
2	21 (2.1%)
3	15 (1.5%)

Query parameter or value	Descriptive statistics
4	8 (0.8%)
5	2 (0.2%)
6	5 (0.5%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	616 (62%)
1	240 (24%)
2	92 (9.2%)
3	24 (2.4%)
4	12 (1.2%)
5	8 (0.8%)
6	4 (0.4%)
7	3 (0.3%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__LOWER	
0	608 (61%)
1	266 (27%)
11	1 (0.1%)
2	73 (7.3%)
3	31 (3.1%)
4	13 (1.3%)
5	4 (0.4%)
6	2 (0.2%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func__UPPER	
0	612 (61%)
1	271 (27%)
12	1 (0.1%)
2	69 (6.9%)
3	20 (2.0%)
4	10 (1.0%)
5	10 (1.0%)
6	5 (0.5%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func__LTRIM	[0.00, 0.00, 12.00] 0.62 / 1.07
SELECT_n_of_non_aggr_func__LOG	
0	844 (84%)
1	118 (12%)
2	26 (2.6%)
3	8 (0.8%)
4	2 (0.2%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func__RTRIM	
0	620 (62%)
1	268 (27%)
2	69 (6.9%)

Query parameter or value	Descriptive statistics
3	22 (2.2%)
4	13 (1.3%)
5	3 (0.3%)
6	4 (0.4%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__ROUND	
0	923 (92%)
1	62 (6.2%)
2	10 (1.0%)
3	5 (0.5%)
SELECT_n_of_non_aggr_func__ABS	
0	919 (92%)
1	68 (6.8%)
2	10 (1.0%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func__DAY	
0	881 (88%)
1	96 (9.6%)
2	16 (1.6%)
3	3 (0.3%)
4	3 (0.3%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__DOW	
0	835 (84%)
1	130 (13%)
2	15 (1.5%)
3	11 (1.1%)
4	6 (0.6%)
5	1 (0.1%)
7	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__MONTH	
0	779 (78%)
1	162 (16%)
12	1 (0.1%)
2	31 (3.1%)
3	13 (1.3%)
4	5 (0.5%)
5	4 (0.4%)
6	3 (0.3%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func__FLOOR	
0	915 (92%)
1	72 (7.2%)
2	11 (1.1%)
3	1 (0.1%)

Query parameter or value	Descriptive statistics
4	1 (0.1%)
SELECT_n_of_non_aggr_func__TRUNC	
0	919 (92%)
1	68 (6.8%)
2	12 (1.2%)
3	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 67.0] 5.0 / 7.6
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 13.00] 0.85 / 1.97
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 13.00] 0.96 / 2.17
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 13.00] 0.89 / 1.99
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 14.00] 0.89 / 2.02
SELECT_n_of_aggr_func__AVG	
0	931 (93%)
1	43 (4.3%)
2	17 (1.7%)
3	7 (0.7%)
4	2 (0.2%)
SELECT_n_of_aggr_func__SUM	
0	931 (93%)
1	45 (4.5%)
2	20 (2.0%)
3	4 (0.4%)
SELECT_n_of_all_aggr_func	[0.0, 3.0, 14.0] 3.8 / 3.2
FROM_n_of_join_paths	
1	154 (15%)
2	148 (15%)
3	168 (17%)
4	194 (19%)
5	165 (17%)
6	171 (17%)
FROM_n_of_super_joins__FULL	
0	552 (55%)
1	312 (31%)
2	104 (10%)
3	29 (2.9%)
4	2 (0.2%)
5	1 (0.1%)
FROM_n_of_super_joins__LEFT	
0	501 (50%)
1	348 (35%)
2	121 (12%)
3	24 (2.4%)
4	6 (0.6%)
FROM_n_of_super_joins__RIGHT	
0	523 (52%)
1	323 (32%)

Query parameter or value	Descriptive statistics
2	130 (13%)
3	21 (2.1%)
4	3 (0.3%)
FROM_n_of_joins__RIGHT	[0.00, 3.00, 14.00] 3.32 / 2.44
FROM_n_of_joins__INNER	[0.00, 3.00, 12.00] 3.41 / 2.49
FROM_n_of_processed_rows	[5, 47,250,086, 220,549,922] 61,087,041 / 49,886,887
WHERE_n_of_predicates	[0.00, 3.00, 12.00] 3.78 / 2.85
WHERE_n_of_attribs_of_type__character_varying	
0	350 (35%)
1	284 (28%)
2	198 (20%)
3	101 (10%)
4	54 (5.4%)
5	10 (1.0%)
6	2 (0.2%)
7	1 (0.1%)
WHERE_n_of_attribs_of_type__integer	
0	375 (38%)
1	302 (30%)
2	174 (17%)
3	95 (9.5%)
4	38 (3.8%)
5	7 (0.7%)
6	9 (0.9%)
WHERE_n_of_attribs_of_type__character	
0	641 (64%)
1	254 (25%)
2	82 (8.2%)
3	20 (2.0%)
4	3 (0.3%)
WHERE_n_of_attribs_of_type__date	
0	801 (80%)
1	160 (16%)
2	33 (3.3%)
3	6 (0.6%)
WHERE_n_of_attribs_of_type__numeric	
0	594 (59%)
1	263 (26%)
2	100 (10%)
3	33 (3.3%)
4	10 (1.0%)
WHERE_n_of_pkey_attribs	
0	544 (54%)
1	293 (29%)
2	114 (11%)
3	43 (4.3%)

Query parameter or value	Descriptive statistics
4	6 (0.6%)
WHERE_n_of_connect_OR	[0.00, 1.00, 11.00] 2.19 / 2.35
WHERE_n_of_operators___between	
0	520 (52%)
1	294 (29%)
2	127 (13%)
3	47 (4.7%)
4	8 (0.8%)
5	2 (0.2%)
6	2 (0.2%)
WHERE_n_of_operators___greater_or_less	
0	325 (33%)
1	291 (29%)
2	188 (19%)
3	106 (11%)
4	53 (5.3%)
5	23 (2.3%)
6	9 (0.9%)
7	4 (0.4%)
8	1 (0.1%)
WHERE_n_of_operators___in	
0	540 (54%)
1	300 (30%)
2	114 (11%)
3	32 (3.2%)
4	13 (1.3%)
5	1 (0.1%)
WHERE_n_of_operators___like	
0	755 (76%)
1	211 (21%)
2	32 (3.2%)
3	2 (0.2%)
WHERE_n_of_non_aggr_func___ABS	
0	957 (96%)
1	43 (4.3%)
WHERE_n_of_non_aggr_func___LOG	
0	934 (93%)
1	62 (6.2%)
2	4 (0.4%)
WHERE_n_of_non_aggr_func___ROUND	
0	959 (96%)
1	38 (3.8%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func___SQRT	
0	884 (88%)
1	109 (11%)

Query parameter or value	Descriptive statistics
2	7 (0.7%)
WHERE_n_of_non_aggr_func__YEAR	
0	971 (97%)
1	27 (2.7%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__FLOOR	
0	961 (96%)
1	38 (3.8%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DOW	
0	971 (97%)
1	28 (2.8%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DAY	
0	967 (97%)
1	31 (3.1%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__TRUNC	
0	964 (96%)
1	35 (3.5%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func__MONTH	
0	968 (97%)
1	30 (3.0%)
2	2 (0.2%)
WHERE_n_of_all_non_aggr_func	
0	631 (63%)
1	275 (28%)
2	73 (7.3%)
3	17 (1.7%)
4	4 (0.4%)
GROUP_BY_n_of_columns	[0.00, 3.00, 12.00] 3.24 / 2.93
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.44 / 2.58
HAVING_n_of_main_predicates__non_scalar_subquery	
0	445 (45%)
1	185 (19%)
2	182 (18%)
3	188 (19%)
HAVING_n_of_main_predicates__scalar_subquery	
0	425 (43%)
1	186 (19%)
2	192 (19%)
3	197 (20%)
HAVING_n_of_subqueries__non_scalar_subquery	
0	445 (45%)
1	185 (19%)

Query parameter or value	Descriptive statistics
<i>2</i>	182 (18%)
<i>3</i>	188 (19%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 20.0] 2.3 / 3.3
HAVING_n_of_processed_rows_by_subqueries	[0, 68,250,376, 2,020,909,534] 207,190,193 / 318,207,758
ORDER_BY_n_of_columns	
<i>0</i>	303 (30%)
<i>1</i>	336 (34%)
<i>2</i>	231 (23%)
<i>3</i>	93 (9.3%)
<i>4</i>	23 (2.3%)
<i>5</i>	8 (0.8%)
<i>6</i>	5 (0.5%)
<i>7</i>	1 (0.1%)
limit	[2, 521, 1,000] 509 / 294
offset	[0, 1, 998] 245 / 323