

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

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

## Scale factor: 5 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, 165] 15 / 22
SELECT_n_of_non_aggr_func__DOW	
0	853 (85%)
1	106 (11%)
2	26 (2.6%)
3	5 (0.5%)
4	5 (0.5%)
5	2 (0.2%)
6	1 (0.1%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func__LOWER	
0	623 (62%)
1	260 (26%)
2	66 (6.6%)
3	30 (3.0%)

Query parameter or value	Descriptive statistics
4	7 (0.7%)
5	6 (0.6%)
6	6 (0.6%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func__LTRIM	[0.00, 0.00, 9.00] 0.58 / 1.01
SELECT_n_of_non_aggr_func__MONTH	
0	808 (81%)
1	140 (14%)
2	31 (3.1%)
3	13 (1.3%)
4	3 (0.3%)
5	2 (0.2%)
6	1 (0.1%)
7	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__ABS	
0	915 (92%)
1	73 (7.3%)
2	8 (0.8%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__LOG	
0	851 (85%)
1	113 (11%)
2	22 (2.2%)
3	8 (0.8%)
4	4 (0.4%)
7	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__ROUND	
0	908 (91%)
1	73 (7.3%)
2	17 (1.7%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__RTRIM	[0.00, 0.00, 11.00] 0.63 / 1.05
SELECT_n_of_non_aggr_func__UPPER	
0	615 (62%)
1	250 (25%)
10	2 (0.2%)
2	81 (8.1%)
3	29 (2.9%)
4	8 (0.8%)
5	7 (0.7%)
6	3 (0.3%)
7	5 (0.5%)

Query parameter or value	Descriptive statistics
SELECT_n_of_non_aggr_func___SQRT	
0	812 (81%)
1	135 (14%)
2	30 (3.0%)
3	14 (1.4%)
4	4 (0.4%)
5	3 (0.3%)
6	2 (0.2%)
SELECT_n_of_non_aggr_func___SUBSTR	[0.00, 0.00, 11.00] 0.58 / 1.08
SELECT_n_of_non_aggr_func___YEAR	
0	822 (82%)
1	131 (13%)
2	32 (3.2%)
3	8 (0.8%)
4	3 (0.3%)
5	1 (0.1%)
6	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___DAY	
0	865 (87%)
1	106 (11%)
2	21 (2.1%)
3	6 (0.6%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___FLOOR	
0	924 (92%)
1	63 (6.3%)
2	10 (1.0%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func___TRUNC	
0	927 (93%)
1	63 (6.3%)
2	7 (0.7%)
3	3 (0.3%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 70.0] 4.8 / 7.5
SELECT_n_of_aggr_func___AVG	
0	928 (93%)
1	52 (5.2%)
2	16 (1.6%)
3	2 (0.2%)
4	2 (0.2%)
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 12.00] 0.83 / 1.94
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 13.00] 0.91 / 2.07
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 11.00] 0.83 / 1.89
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 11.00] 0.92 / 1.94

Query parameter or value	Descriptive statistics
SELECT_n_of_aggr_func__SUM	
0	931 (93%)
1	46 (4.6%)
2	16 (1.6%)
3	5 (0.5%)
4	2 (0.2%)
SELECT_n_of_all_aggr_func	[0.0, 3.0, 13.0] 3.7 / 3.1
FROM_n_of_join_paths	
1	156 (16%)
2	182 (18%)
3	169 (17%)
4	165 (17%)
5	157 (16%)
6	171 (17%)
FROM_n_of_super_joins__FULL	
0	547 (55%)
1	311 (31%)
2	108 (11%)
3	32 (3.2%)
4	2 (0.2%)
FROM_n_of_super_joins__RIGHT	
0	558 (56%)
1	310 (31%)
2	100 (10%)
3	29 (2.9%)
4	3 (0.3%)
FROM_n_of_super_joins__LEFT	
0	551 (55%)
1	300 (30%)
2	122 (12%)
3	22 (2.2%)
4	5 (0.5%)
FROM_n_of_joins__INNER	[0.00, 3.00, 12.00] 3.32 / 2.47
FROM_n_of_joins__RIGHT	[0.00, 3.00, 15.00] 3.16 / 2.44
FROM_n_of_processed_rows	[5, 46,500,033, 225,249,768] 56,495,202 / 44,842,961
WHERE_n_of_predicates	[0.00, 3.00, 13.00] 3.81 / 2.94
WHERE_n_of_attribs_of_type__character_varying	
0	376 (38%)
1	263 (26%)
2	192 (19%)
3	98 (9.8%)
4	50 (5.0%)
5	16 (1.6%)
6	5 (0.5%)
WHERE_n_of_attribs_of_type__date	
0	793 (79%)

Query parameter or value	Descriptive statistics
1	180 (18%)
2	23 (2.3%)
3	4 (0.4%)
WHERE_n_of_attribs_of_type__integer	
0	361 (36%)
1	305 (31%)
2	181 (18%)
3	94 (9.4%)
4	43 (4.3%)
5	10 (1.0%)
6	6 (0.6%)
WHERE_n_of_attribs_of_type__numeric	
0	564 (56%)
1	287 (29%)
2	101 (10%)
3	36 (3.6%)
4	9 (0.9%)
5	3 (0.3%)
WHERE_n_of_attribs_of_type__character	
0	648 (65%)
1	261 (26%)
2	73 (7.3%)
3	16 (1.6%)
4	2 (0.2%)
WHERE_n_of_pkey_attribs	
0	549 (55%)
1	285 (29%)
2	127 (13%)
3	25 (2.5%)
4	14 (1.4%)
WHERE_n_of_connect_OR	[0.00, 1.50, 12.00] 2.17 / 2.40
WHERE_n_of_operators__between	
0	556 (56%)
1	267 (27%)
2	124 (12%)
3	37 (3.7%)
4	14 (1.4%)
5	1 (0.1%)
6	1 (0.1%)
WHERE_n_of_operators__greater_or_less	[0.00, 1.00, 9.00] 1.42 / 1.46
WHERE_n_of_operators__in	
0	538 (54%)
1	291 (29%)
2	123 (12%)
3	38 (3.8%)
4	8 (0.8%)

Query parameter or value	Descriptive statistics
5	2 (0.2%)
WHERE_n_of_operators_like	
0	793 (79%)
1	177 (18%)
2	23 (2.3%)
3	6 (0.6%)
4	1 (0.1%)
WHERE_n_of_non_aggr_func_DOW	
0	963 (96%)
1	36 (3.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_LOG	
0	924 (92%)
1	72 (7.2%)
2	4 (0.4%)
WHERE_n_of_non_aggr_func_SQRT	
0	875 (88%)
1	113 (11%)
2	11 (1.1%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func_ABS	
0	952 (95%)
1	47 (4.7%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_TRUNC	
0	962 (96%)
1	37 (3.7%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_YEAR	
0	969 (97%)
1	31 (3.1%)
WHERE_n_of_non_aggr_func_ROUND	
0	965 (97%)
1	35 (3.5%)
WHERE_n_of_non_aggr_func_FLOOR	
0	959 (96%)
1	41 (4.1%)
WHERE_n_of_non_aggr_func_DAY	
0	972 (97%)
1	28 (2.8%)
WHERE_n_of_non_aggr_func_MONTH	
0	970 (97%)
1	29 (2.9%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	631 (63%)

Query parameter or value	Descriptive statistics
<i>1</i>	268 (27%)
<i>2</i>	69 (6.9%)
<i>3</i>	25 (2.5%)
<i>4</i>	6 (0.6%)
<i>5</i>	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 12.00] 3.14 / 2.86
HAVING_n_of_main_predicates	
<i>0</i>	427 (43%)
<i>1</i>	196 (20%)
<i>2</i>	187 (19%)
<i>3</i>	190 (19%)
ORDER_BY_n_of_columns	
<i>0</i>	321 (32%)
<i>1</i>	322 (32%)
<i>2</i>	238 (24%)
<i>3</i>	85 (8.5%)
<i>4</i>	22 (2.2%)
<i>5</i>	7 (0.7%)
<i>6</i>	3 (0.3%)
<i>7</i>	2 (0.2%)
limit	[2, 521, 999] 503 / 287
offset	[0, 0, 996] 239 / 321