

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

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

## Scale factor: 1000 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, 8, 170] 14 / 21
SELECT_n_of_non_aggr_func__SUBSTR	
0	630 (63%)
1	248 (25%)
2	77 (7.7%)
3	25 (2.5%)
4	7 (0.7%)
5	6 (0.6%)
6	6 (0.6%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__FLOOR	
0	932 (93%)
1	59 (5.9%)
2	7 (0.7%)
3	1 (0.1%)

Query parameter or value	Descriptive statistics
4	1 (0.1%)
SELECT_n_of_non_aggr_func__LOWER	
0	609 (61%)
1	265 (27%)
2	89 (8.9%)
3	19 (1.9%)
4	6 (0.6%)
5	8 (0.8%)
6	1 (0.1%)
7	1 (0.1%)
8	2 (0.2%)
SELECT_n_of_non_aggr_func__LTRIM	
0	623 (62%)
1	267 (27%)
2	78 (7.8%)
3	18 (1.8%)
4	10 (1.0%)
5	2 (0.2%)
6	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__RTRIM	
0	617 (62%)
1	265 (27%)
2	70 (7.0%)
3	30 (3.0%)
4	7 (0.7%)
5	5 (0.5%)
6	4 (0.4%)
7	1 (0.1%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func__UPPER	[0.00, 0.00, 11.00] 0.58 / 0.98
SELECT_n_of_non_aggr_func__TRUNC	
0	923 (92%)
1	65 (6.5%)
2	10 (1.0%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__LOG	
0	865 (87%)
1	105 (11%)
2	24 (2.4%)
3	5 (0.5%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__MONTH	
0	801 (80%)
1	140 (14%)

Query parameter or value	Descriptive statistics
2	34 (3.4%)
3	12 (1.2%)
4	5 (0.5%)
5	4 (0.4%)
6	2 (0.2%)
7	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___ABS	
0	918 (92%)
1	73 (7.3%)
2	9 (0.9%)
SELECT_n_of_non_aggr_func___DOW	
0	857 (86%)
1	107 (11%)
2	28 (2.8%)
3	3 (0.3%)
4	2 (0.2%)
5	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___ROUND	
0	923 (92%)
1	62 (6.2%)
2	12 (1.2%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func___YEAR	
0	831 (83%)
1	122 (12%)
2	27 (2.7%)
3	9 (0.9%)
4	5 (0.5%)
5	4 (0.4%)
7	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___DAY	
0	900 (90%)
1	77 (7.7%)
2	11 (1.1%)
3	7 (0.7%)
4	3 (0.3%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___SQRT	
0	809 (81%)
1	143 (14%)
2	26 (2.6%)
3	8 (0.8%)

Query parameter or value	Descriptive statistics
4	7 (0.7%)
5	4 (0.4%)
6	1 (0.1%)
7	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 67.0] 4.6 / 6.5
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 15.00] 0.88 / 1.95
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 13.00] 0.92 / 2.11
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 12.00] 0.86 / 1.96
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 12.00] 0.80 / 1.83
SELECT_n_of_aggr_func__AVG	
0	922 (92%)
1	52 (5.2%)
2	22 (2.2%)
3	4 (0.4%)
SELECT_n_of_aggr_func__SUM	
0	932 (93%)
1	41 (4.1%)
2	19 (1.9%)
3	7 (0.7%)
4	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.00, 3.00, 15.00] 3.66 / 3.11
FROM_n_of_join_paths	
1	168 (17%)
2	177 (18%)
3	177 (18%)
4	150 (15%)
5	174 (17%)
6	154 (15%)
FROM_n_of_super_joins__FULL	
0	569 (57%)
1	296 (30%)
2	105 (11%)
3	29 (2.9%)
4	1 (0.1%)
FROM_n_of_super_joins__LEFT	
0	554 (55%)
1	316 (32%)
2	101 (10%)
3	26 (2.6%)
4	3 (0.3%)
FROM_n_of_super_joins__RIGHT	
0	549 (55%)
1	318 (32%)
2	106 (11%)
3	24 (2.4%)

Query parameter or value	Descriptive statistics
4	3 (0.3%)
FROM_n_of_joins__INNER	[0.00, 3.00, 14.00] 3.23 / 2.54
FROM_n_of_joins__RIGHT	[0.00, 3.00, 15.00] 3.21 / 2.52
FROM_n_of_processed_rows	[5, 9,299,989,709, 44,069,938,339] 11,820,885,031 / 9,734,684,306
WHERE_n_of_predicates	[0.00, 3.00, 12.00] 3.77 / 2.86
WHERE_n_of_attribs_of_type__character	
0	673 (67%)
1	242 (24%)
2	63 (6.3%)
3	18 (1.8%)
4	3 (0.3%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type__character_varying	
0	371 (37%)
1	297 (30%)
2	166 (17%)
3	105 (11%)
4	47 (4.7%)
5	11 (1.1%)
6	3 (0.3%)
WHERE_n_of_attribs_of_type__numeric	
0	569 (57%)
1	283 (28%)
2	113 (11%)
3	27 (2.7%)
4	8 (0.8%)
WHERE_n_of_attribs_of_type__date	
0	777 (78%)
1	178 (18%)
2	38 (3.8%)
3	6 (0.6%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type__integer	
0	334 (33%)
1	337 (34%)
2	181 (18%)
3	89 (8.9%)
4	40 (4.0%)
5	16 (1.6%)
6	2 (0.2%)
7	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	515 (52%)
1	321 (32%)
2	113 (11%)

Query parameter or value	Descriptive statistics
3	41 (4.1%)
4	9 (0.9%)
5	1 (0.1%)
WHERE_n_of_connect_OR	[0.00, 1.00, 11.00] 2.12 / 2.32
WHERE_n_of_operators___greater_or_less	
0	316 (32%)
1	292 (29%)
2	193 (19%)
3	108 (11%)
4	51 (5.1%)
5	29 (2.9%)
6	7 (0.7%)
7	4 (0.4%)
WHERE_n_of_operators___in	
0	515 (52%)
1	299 (30%)
2	137 (14%)
3	40 (4.0%)
4	7 (0.7%)
5	2 (0.2%)
WHERE_n_of_operators___between	
0	553 (55%)
1	286 (29%)
2	115 (12%)
3	34 (3.4%)
4	12 (1.2%)
WHERE_n_of_operators___like	
0	804 (80%)
1	171 (17%)
2	21 (2.1%)
3	4 (0.4%)
WHERE_n_of_non_aggr_func___YEAR	
0	955 (96%)
1	43 (4.3%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___SQRT	
0	864 (86%)
1	131 (13%)
2	5 (0.5%)
WHERE_n_of_non_aggr_func___ABS	
0	958 (96%)
1	40 (4.0%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___FLOOR	
0	965 (97%)
1	35 (3.5%)

Query parameter or value	Descriptive statistics
WHERE_n_of_non_aggr_func__TRUNC	
0	960 (96%)
1	37 (3.7%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func__DOW	
0	967 (97%)
1	31 (3.1%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__ROUND	
0	953 (95%)
1	46 (4.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DAY	
0	967 (97%)
1	33 (3.3%)
WHERE_n_of_non_aggr_func__MONTH	
0	958 (96%)
1	39 (3.9%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func__LOG	
0	922 (92%)
1	75 (7.5%)
2	3 (0.3%)
WHERE_n_of_all_non_aggr_func	
0	609 (61%)
1	265 (27%)
2	96 (9.6%)
3	26 (2.6%)
4	3 (0.3%)
5	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 3.00, 12.00] 3.15 / 2.86
HAVING_n_of_main_predicates	[0.00, 3.00, 9.00] 3.35 / 2.58
HAVING_n_of_main_predicates__non_scalar_subquery	
0	446 (45%)
1	185 (19%)
2	178 (18%)
3	191 (19%)
HAVING_n_of_main_predicates__scalar_subquery	
0	438 (44%)
1	180 (18%)
2	184 (18%)
3	198 (20%)
HAVING_n_of_subqueries__non_scalar_subquery	
0	446 (45%)
1	185 (19%)
2	178 (18%)

Query parameter or value	Descriptive statistics
<i>3</i>	191 (19%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 23.0] 2.4 / 3.6
HAVING_n_of_processed_rows_by_subquery	[0, 13, 519, 990, 129, 644, 999, 179, 620] 40,469,068,841 / 69,017,317,109
ORDER_BY_n_of_columns	
<i>0</i>	301 (30%)
<i>1</i>	333 (33%)
<i>2</i>	239 (24%)
<i>3</i>	97 (9.7%)
<i>4</i>	21 (2.1%)
<i>5</i>	3 (0.3%)
<i>6</i>	4 (0.4%)
<i>7</i>	2 (0.2%)
limit	[3, 506, 1,000] 496 / 283
offset	[0, 0, 998] 237 / 321