

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

```
## Scenario: 1-3jp_rjpp__g__h0-3ss_h0nss

## Scale factor: 100 GB

## The 1000-query pack was generated on: 2023-11-07

## 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, 6, 100] 10 / 12
SELECT_n_of_non_aggr_func__DAY	
0	928 (93%)
1	66 (6.6%)
2	5 (0.5%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__LOWER	
0	699 (70%)
1	227 (23%)
2	58 (5.8%)
3	9 (0.9%)
4	5 (0.5%)
6	2 (0.2%)
SELECT_n_of_non_aggr_func__MONTH	
0	858 (86%)

Query parameter or value	Descriptive statistics
1	114 (11%)
2	16 (1.6%)
3	5 (0.5%)
4	4 (0.4%)
6	3 (0.3%)
SELECT_n_of_non_aggr_func__ROUND	
0	952 (95%)
1	45 (4.5%)
2	3 (0.3%)
SELECT_n_of_non_aggr_func__RTRIM	
0	685 (69%)
1	263 (26%)
2	37 (3.7%)
3	9 (0.9%)
4	4 (0.4%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	705 (71%)
1	236 (24%)
2	46 (4.6%)
3	10 (1.0%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__UPPER	
0	688 (69%)
1	235 (24%)
2	61 (6.1%)
3	10 (1.0%)
4	4 (0.4%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__YEAR	
0	876 (88%)
1	106 (11%)
2	12 (1.2%)
3	4 (0.4%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__SQRT	
0	862 (86%)
1	109 (11%)
2	21 (2.1%)
3	3 (0.3%)
4	3 (0.3%)
5	1 (0.1%)

Query parameter or value	Descriptive statistics
6	1 (0.1%)
SELECT_n_of_non_aggr_func___LTRIM	
0	707 (71%)
1	226 (23%)
2	51 (5.1%)
3	11 (1.1%)
4	4 (0.4%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___TRUNC	
0	949 (95%)
1	50 (5.0%)
2	1 (0.1%)
SELECT_n_of_non_aggr_func___LOG	
0	902 (90%)
1	86 (8.6%)
2	7 (0.7%)
3	4 (0.4%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___ABS	
0	946 (95%)
1	51 (5.1%)
2	3 (0.3%)
SELECT_n_of_non_aggr_func___DOW	
0	907 (91%)
1	73 (7.3%)
2	14 (1.4%)
3	4 (0.4%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___FLOOR	
0	957 (96%)
1	41 (4.1%)
2	2 (0.2%)
SELECT_n_of_all_non_aggr_func	[0.00, 2.00, 44.00] 3.02 / 3.92
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 9.00] 0.63 / 1.40
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 10.00] 0.64 / 1.45
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 10.00] 0.65 / 1.51
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 10.00] 0.62 / 1.48
SELECT_n_of_aggr_func___AVG	
0	936 (94%)
1	46 (4.6%)
2	13 (1.3%)
3	5 (0.5%)
SELECT_n_of_aggr_func___SUM	
0	940 (94%)
1	46 (4.6%)
2	12 (1.2%)

Query parameter or value	Descriptive statistics
3	2 (0.2%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 12.00] 2.71 / 2.31
FROM_n_of_join_paths	
1	319 (32%)
2	349 (35%)
3	332 (33%)
FROM_n_of_super_joins___FULL	
0	753 (75%)
1	231 (23%)
2	16 (1.6%)
FROM_n_of_super_joins___LEFT	
0	762 (76%)
1	211 (21%)
2	27 (2.7%)
FROM_n_of_super_joins___RIGHT	
0	767 (77%)
1	218 (22%)
2	15 (1.5%)
FROM_n_of_joins___INNER	
0	237 (24%)
1	263 (26%)
2	218 (22%)
3	137 (14%)
4	89 (8.9%)
5	41 (4.1%)
6	11 (1.1%)
7	2 (0.2%)
8	2 (0.2%)
FROM_n_of_joins___RIGHT	
0	244 (24%)
1	263 (26%)
2	201 (20%)
3	150 (15%)
4	85 (8.5%)
5	39 (3.9%)
6	14 (1.4%)
7	2 (0.2%)
8	2 (0.2%)
FROM_n_of_processed_rows	[5, 681,037,902, 2,295,113,756] 630,994,140 / 596,369,638
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.79 / 2.21
WHERE_n_of_attribs_of_type___character	
0	717 (72%)
1	224 (22%)
2	44 (4.4%)
3	14 (1.4%)

Query parameter or value	Descriptive statistics
4	1 (0.1%)
WHERE_n_of_attribs_of_type__character_varying	
0	457 (46%)
1	310 (31%)
2	154 (15%)
3	60 (6.0%)
4	16 (1.6%)
5	2 (0.2%)
6	1 (0.1%)
WHERE_n_of_attribs_of_type__integer	
0	445 (45%)
1	329 (33%)
2	146 (15%)
3	63 (6.3%)
4	14 (1.4%)
5	2 (0.2%)
6	1 (0.1%)
WHERE_n_of_attribs_of_type__numeric	
0	621 (62%)
1	280 (28%)
2	86 (8.6%)
3	12 (1.2%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	835 (84%)
1	150 (15%)
2	14 (1.4%)
3	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	612 (61%)
1	296 (30%)
2	77 (7.7%)
3	14 (1.4%)
5	1 (0.1%)
WHERE_n_of_connect_OR	
0	447 (45%)
1	198 (20%)
2	128 (13%)
3	92 (9.2%)
4	71 (7.1%)
5	42 (4.2%)
6	16 (1.6%)
7	5 (0.5%)
8	1 (0.1%)
WHERE_n_of_operators__between	
0	602 (60%)

Query parameter or value	Descriptive statistics
1	276 (28%)
2	96 (9.6%)
3	21 (2.1%)
4	4 (0.4%)
5	1 (0.1%)
WHERE_n_of_operators___greater_or_less	
0	439 (44%)
1	292 (29%)
2	155 (16%)
3	84 (8.4%)
4	21 (2.1%)
5	7 (0.7%)
6	2 (0.2%)
WHERE_n_of_operators___in	
0	599 (60%)
1	297 (30%)
2	77 (7.7%)
3	24 (2.4%)
4	3 (0.3%)
WHERE_n_of_operators___like	
0	827 (83%)
1	159 (16%)
2	13 (1.3%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func___DAY	
0	979 (98%)
1	21 (2.1%)
WHERE_n_of_non_aggr_func___LOG	
0	945 (95%)
1	53 (5.3%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___SQRT	
0	905 (91%)
1	88 (8.8%)
2	7 (0.7%)
WHERE_n_of_non_aggr_func___ABS	
0	969 (97%)
1	31 (3.1%)
WHERE_n_of_non_aggr_func___MONTH	
0	972 (97%)
1	28 (2.8%)
WHERE_n_of_non_aggr_func___TRUNC	
0	962 (96%)
1	38 (3.8%)
WHERE_n_of_non_aggr_func___ROUND	
0	969 (97%)

Query parameter or value	Descriptive statistics
1	31 (3.1%)
WHERE_n_of_non_aggr_func__YEAR	
0	971 (97%)
1	28 (2.8%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DOW	
0	983 (98%)
1	17 (1.7%)
WHERE_n_of_non_aggr_func__FLOOR	
0	971 (97%)
1	29 (2.9%)
WHERE_n_of_all_non_aggr_func	
0	709 (71%)
1	215 (22%)
2	63 (6.3%)
3	11 (1.1%)
5	2 (0.2%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.26 / 2.08
HAVING_n_of_main_predicates	
0	281 (28%)
1	98 (9.8%)
2	164 (16%)
3	186 (19%)
4	133 (13%)
5	87 (8.7%)
6	51 (5.1%)
HAVING_n_of_main_predicates__scalar_subquery	
0	427 (43%)
1	196 (20%)
2	201 (20%)
3	176 (18%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 24.0] 2.3 / 3.6
HAVING_n_of_processed_rows_by_subqueries	[0, 45,000,083, 36,601,819,296] 1,592,277,252 / 3,821,839,071
ORDER_BY_n_of_columns	
0	296 (30%)
1	395 (40%)
2	228 (23%)
3	57 (5.7%)
4	11 (1.1%)
5	7 (0.7%)
6	4 (0.4%)
7	2 (0.2%)
limit	[1, 513, 1,000] 506 / 293
offset	[0, 3, 999] 262 / 332