

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

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

## Scale factor: 200 GB

## The 1000-query pack was generated on: 2023-11-13

## 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, 186] 15 / 23
SELECT_n_of_non_aggr_func__RTRIM	[0.00, 0.00, 9.00] 0.64 / 1.18
SELECT_n_of_non_aggr_func__SQRT	
0	815 (82%)
1	131 (13%)
2	23 (2.3%)
3	15 (1.5%)
4	8 (0.8%)
5	4 (0.4%)
6	3 (0.3%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__UPPER	[0.00, 0.00, 10.00] 0.61 / 1.10
SELECT_n_of_non_aggr_func__FLOOR	
0	916 (92%)
1	68 (6.8%)

Query parameter or value	Descriptive statistics
2	11 (1.1%)
3	3 (0.3%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__TRUNC	
0	917 (92%)
1	68 (6.8%)
2	13 (1.3%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__DAY	
0	888 (89%)
1	89 (8.9%)
2	16 (1.6%)
3	3 (0.3%)
4	2 (0.2%)
6	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__DOW	
0	844 (84%)
1	107 (11%)
2	33 (3.3%)
3	9 (0.9%)
4	6 (0.6%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__MONTH	[0.00, 0.00, 11.00] 0.33 / 0.95
SELECT_n_of_non_aggr_func__SUBSTR	[0.00, 0.00, 10.00] 0.61 / 1.10
SELECT_n_of_non_aggr_func__YEAR	
0	820 (82%)
1	125 (13%)
2	29 (2.9%)
3	10 (1.0%)
4	7 (0.7%)
5	6 (0.6%)
7	2 (0.2%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__LTRIM	
0	599 (60%)
1	283 (28%)
2	76 (7.6%)
3	19 (1.9%)
4	7 (0.7%)
5	9 (0.9%)
6	2 (0.2%)
8	2 (0.2%)
9	3 (0.3%)
SELECT_n_of_non_aggr_func__ABS	

Query parameter or value	Descriptive statistics
0	925 (93%)
1	65 (6.5%)
2	7 (0.7%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func__LOG	
0	861 (86%)
1	105 (11%)
2	21 (2.1%)
3	5 (0.5%)
4	6 (0.6%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func__LOWER	
0	652 (65%)
1	226 (23%)
2	69 (6.9%)
3	23 (2.3%)
4	11 (1.1%)
5	7 (0.7%)
6	7 (0.7%)
7	4 (0.4%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__ROUND	
0	907 (91%)
1	83 (8.3%)
2	5 (0.5%)
3	4 (0.4%)
4	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 72.0] 5.0 / 8.1
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 11.00] 0.75 / 1.86
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 13.00] 0.84 / 1.89
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 13.00] 0.89 / 1.96
SELECT_n_of_aggr_func__AVG	
0	947 (95%)
1	39 (3.9%)
2	10 (1.0%)
3	4 (0.4%)
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 13.00] 0.86 / 1.99
SELECT_n_of_aggr_func__SUM	
0	929 (93%)
1	48 (4.8%)
2	20 (2.0%)
3	3 (0.3%)
SELECT_n_of_all_aggr_func	[0.0, 3.0, 13.0] 3.5 / 3.1
FROM_n_of_join_paths	
1	186 (19%)
2	162 (16%)

Query parameter or value	Descriptive statistics
3	160 (16%)
4	145 (15%)
5	176 (18%)
6	171 (17%)
FROM_n_of_super_joins___FULL	
0	572 (57%)
1	298 (30%)
2	99 (9.9%)
3	30 (3.0%)
4	1 (0.1%)
FROM_n_of_super_joins___LEFT	
0	569 (57%)
1	292 (29%)
2	109 (11%)
3	26 (2.6%)
4	4 (0.4%)
FROM_n_of_super_joins___RIGHT	
0	537 (54%)
1	310 (31%)
2	118 (12%)
3	32 (3.2%)
4	3 (0.3%)
FROM_n_of_joins___INNER	[0.00, 3.00, 12.00] 3.20 / 2.46
FROM_n_of_joins___RIGHT	[0.00, 3.00, 13.00] 3.33 / 2.58
FROM_n_of_processed_rows	[5, 1,854,018,502, 8,552,110,659] 2,291,296,290 / 1,942,409,868
WHERE_n_of_predicates	[0.0, 3.0, 14.0] 3.9 / 3.0
WHERE_n_of_attribs_of_type___integer	
0	363 (36%)
1	300 (30%)
2	184 (18%)
3	91 (9.1%)
4	42 (4.2%)
5	15 (1.5%)
6	5 (0.5%)
WHERE_n_of_attribs_of_type___character	
0	653 (65%)
1	247 (25%)
2	73 (7.3%)
3	23 (2.3%)
4	4 (0.4%)
WHERE_n_of_attribs_of_type___character_varying	
0	360 (36%)
1	277 (28%)
2	184 (18%)
3	101 (10%)

Query parameter or value	Descriptive statistics
4	52 (5.2%)
5	24 (2.4%)
6	2 (0.2%)
WHERE_n_of_attribs_of_type__date	
0	794 (79%)
1	164 (16%)
2	39 (3.9%)
3	3 (0.3%)
WHERE_n_of_attribs_of_type__numeric	
0	573 (57%)
1	283 (28%)
2	93 (9.3%)
3	35 (3.5%)
4	15 (1.5%)
6	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	523 (52%)
1	321 (32%)
2	103 (10%)
3	43 (4.3%)
4	9 (0.9%)
5	1 (0.1%)
WHERE_n_of_connect_OR	[0.00, 1.00, 13.00] 2.26 / 2.54
WHERE_n_of_operators__greater_or_less	
0	327 (33%)
1	272 (27%)
2	185 (19%)
3	108 (11%)
4	58 (5.8%)
5	29 (2.9%)
6	16 (1.6%)
7	5 (0.5%)
WHERE_n_of_operators__in	
0	551 (55%)
1	263 (26%)
2	130 (13%)
3	39 (3.9%)
4	11 (1.1%)
5	5 (0.5%)
6	1 (0.1%)
WHERE_n_of_operators__like	
0	794 (79%)
1	172 (17%)
2	27 (2.7%)
3	4 (0.4%)
4	3 (0.3%)

Query parameter or value	Descriptive statistics
WHERE_n_of_operators__between	
0	534 (53%)
1	283 (28%)
2	121 (12%)
3	48 (4.8%)
4	11 (1.1%)
5	3 (0.3%)
WHERE_n_of_non_aggr_func__DOW	
0	968 (97%)
1	30 (3.0%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__FLOOR	
0	955 (96%)
1	41 (4.1%)
2	4 (0.4%)
WHERE_n_of_non_aggr_func__SQRT	
0	887 (89%)
1	109 (11%)
2	4 (0.4%)
WHERE_n_of_non_aggr_func__TRUNC	
0	959 (96%)
1	39 (3.9%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__MONTH	
0	965 (97%)
1	35 (3.5%)
WHERE_n_of_non_aggr_func__LOG	
0	927 (93%)
1	68 (6.8%)
2	5 (0.5%)
WHERE_n_of_non_aggr_func__ROUND	
0	946 (95%)
1	54 (5.4%)
WHERE_n_of_non_aggr_func__ABS	
0	962 (96%)
1	35 (3.5%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func__DAY	
0	961 (96%)
1	38 (3.8%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__YEAR	
0	958 (96%)
1	39 (3.9%)
2	3 (0.3%)
WHERE_n_of_all_non_aggr_func	

Query parameter or value	Descriptive statistics
0	635 (64%)
1	240 (24%)
2	88 (8.8%)
3	30 (3.0%)
4	6 (0.6%)
6	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 12.00] 3.12 / 2.91
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.38 / 2.59
HAVING_n_of_main_predicates__non_scalar_subquery	
0	420 (42%)
1	183 (18%)
2	208 (21%)
3	189 (19%)
HAVING_n_of_main_predicates__scalar_subquery	
0	447 (45%)
1	160 (16%)
2	189 (19%)
3	204 (20%)
HAVING_n_of_subqueries__non_scalar_subquery	
0	420 (42%)
1	183 (18%)
2	208 (21%)
3	189 (19%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 20.0] 2.4 / 3.5
HAVING_n_of_processed_rows_by_subqueries	[0, 2, 101,000,245, 122,551,402,029] 8,217,034,412 / 14,113,262,515
ORDER_BY_n_of_columns	
0	302 (30%)
1	349 (35%)
2	230 (23%)
3	90 (9.0%)
4	14 (1.4%)
5	11 (1.1%)
6	2 (0.2%)
7	1 (0.1%)
8	1 (0.1%)
limit	[1, 502, 999] 500 / 288
offset	[0, 0, 999] 253 / 332