

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

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

## Scale factor: 500 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, 8, 176] 16 / 25
SELECT_n_of_non_aggr_func__UPPER	
0	633 (63%)
1	240 (24%)
2	83 (8.3%)
3	21 (2.1%)
4	9 (0.9%)
5	10 (1.0%)
6	1 (0.1%)
7	3 (0.3%)
SELECT_n_of_non_aggr_func__LTRIM	
0	626 (63%)
1	247 (25%)
2	89 (8.9%)
3	16 (1.6%)

Query parameter or value	Descriptive statistics
4	6 (0.6%)
5	10 (1.0%)
6	2 (0.2%)
7	4 (0.4%)
SELECT__n_of__non_aggr_func__DAY	
0	894 (89%)
1	86 (8.6%)
2	12 (1.2%)
3	6 (0.6%)
4	2 (0.2%)
SELECT__n_of__non_aggr_func__DOW	
0	834 (83%)
1	124 (12%)
10	1 (0.1%)
2	30 (3.0%)
3	3 (0.3%)
4	7 (0.7%)
6	1 (0.1%)
SELECT__n_of__non_aggr_func__MONTH	
0	814 (81%)
1	122 (12%)
13	1 (0.1%)
2	34 (3.4%)
3	11 (1.1%)
4	6 (0.6%)
5	7 (0.7%)
6	4 (0.4%)
8	1 (0.1%)
SELECT__n_of__non_aggr_func__YEAR	
0	823 (82%)
1	124 (12%)
2	33 (3.3%)
3	5 (0.5%)
4	10 (1.0%)
5	2 (0.2%)
6	1 (0.1%)
7	1 (0.1%)
8	1 (0.1%)
SELECT__n_of__non_aggr_func__ABS	
0	920 (92%)
1	67 (6.7%)
2	12 (1.2%)
3	1 (0.1%)
SELECT__n_of__non_aggr_func__LOWER	
0	643 (64%)
1	231 (23%)

Query parameter or value	Descriptive statistics
2	72 (7.2%)
3	28 (2.8%)
4	14 (1.4%)
5	7 (0.7%)
6	4 (0.4%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__RTRIM	
0	647 (65%)
1	239 (24%)
2	70 (7.0%)
3	27 (2.7%)
4	7 (0.7%)
5	5 (0.5%)
6	3 (0.3%)
7	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__SQRT	
0	833 (83%)
1	119 (12%)
2	27 (2.7%)
3	13 (1.3%)
4	2 (0.2%)
5	5 (0.5%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	643 (64%)
1	258 (26%)
2	61 (6.1%)
3	19 (1.9%)
4	8 (0.8%)
5	4 (0.4%)
6	2 (0.2%)
7	3 (0.3%)
8	2 (0.2%)
SELECT_n_of_non_aggr_func__LOG	
0	879 (88%)
1	104 (10%)
2	9 (0.9%)
3	5 (0.5%)
4	3 (0.3%)
SELECT_n_of_non_aggr_func__ROUND	
0	923 (92%)
1	65 (6.5%)
2	11 (1.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__TRUNC	

Query parameter or value	Descriptive statistics
0	933 (93%)
1	55 (5.5%)
2	11 (1.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__FLOOR	
0	925 (93%)
1	63 (6.3%)
2	10 (1.0%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 69.0] 4.6 / 7.0
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 12.00] 0.81 / 1.84
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 13.00] 0.84 / 1.94
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 12.00] 0.90 / 2.04
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 12.00] 0.77 / 1.87
SELECT_n_of_aggr_func__SUM	
0	945 (95%)
1	39 (3.9%)
2	11 (1.1%)
3	5 (0.5%)
SELECT_n_of_aggr_func__AVG	
0	921 (92%)
1	60 (6.0%)
2	9 (0.9%)
3	7 (0.7%)
4	3 (0.3%)
SELECT_n_of_all_aggr_func	[0.00, 3.00, 14.00] 3.50 / 3.10
FROM_n_of_join_paths	
1	188 (19%)
2	172 (17%)
3	149 (15%)
4	148 (15%)
5	177 (18%)
6	166 (17%)
FROM_n_of_super_joins__FULL	
0	566 (57%)
1	305 (31%)
2	93 (9.3%)
3	34 (3.4%)
4	2 (0.2%)
FROM_n_of_super_joins__RIGHT	
0	546 (55%)
1	291 (29%)
2	132 (13%)
3	25 (2.5%)
4	5 (0.5%)

Query parameter or value	Descriptive statistics
5	1 (0.1%)
FROM_n_of_super_joins__LEFT	
0	568 (57%)
1	298 (30%)
2	107 (11%)
3	20 (2.0%)
4	6 (0.6%)
5	1 (0.1%)
FROM_n_of_joins__INNER	[0.00, 3.00, 13.00] 3.20 / 2.44
FROM_n_of_joins__RIGHT	[0.00, 3.00, 14.00] 3.18 / 2.50
FROM_n_of_processed_rows	[5, 4,650,028,272, 22,530,169,477] 5,956,351,882 / 5,048,146,665
WHERE_n_of_predicates	[0.00, 3.00, 13.00] 3.75 / 3.04
WHERE_n_of_attribs_of_type__character	
0	672 (67%)
1	241 (24%)
2	65 (6.5%)
3	17 (1.7%)
4	5 (0.5%)
WHERE_n_of_attribs_of_type__character_varying	
0	372 (37%)
1	290 (29%)
2	179 (18%)
3	102 (10%)
4	35 (3.5%)
5	16 (1.6%)
6	5 (0.5%)
8	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	769 (77%)
1	180 (18%)
2	40 (4.0%)
3	8 (0.8%)
4	3 (0.3%)
WHERE_n_of_attribs_of_type__integer	
0	386 (39%)
1	286 (29%)
2	181 (18%)
3	98 (9.8%)
4	31 (3.1%)
5	11 (1.1%)
6	6 (0.6%)
7	1 (0.1%)
WHERE_n_of_attribs_of_type__numeric	
0	582 (58%)
1	265 (27%)

Query parameter or value	Descriptive statistics
2	101 (10%)
3	41 (4.1%)
4	10 (1.0%)
5	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	559 (56%)
1	302 (30%)
2	103 (10%)
3	28 (2.8%)
4	6 (0.6%)
5	2 (0.2%)
WHERE_n_of_connect_OR	[0.00, 1.00, 11.00] 2.16 / 2.49
WHERE_n_of_operators__between	
0	540 (54%)
1	284 (28%)
2	118 (12%)
3	44 (4.4%)
4	12 (1.2%)
5	2 (0.2%)
WHERE_n_of_operators__in	
0	550 (55%)
1	295 (30%)
2	104 (10%)
3	40 (4.0%)
4	9 (0.9%)
5	2 (0.2%)
WHERE_n_of_operators__greater_or_less	
0	362 (36%)
1	262 (26%)
2	176 (18%)
3	103 (10%)
4	56 (5.6%)
5	24 (2.4%)
6	13 (1.3%)
7	3 (0.3%)
8	1 (0.1%)
WHERE_n_of_operators__like	
0	782 (78%)
1	187 (19%)
2	25 (2.5%)
3	5 (0.5%)
4	1 (0.1%)
WHERE_n_of_non_aggr_func__MONTH	
0	960 (96%)
1	38 (3.8%)
2	2 (0.2%)

Query parameter or value	Descriptive statistics
WHERE_n_of_non_aggr_func__SQRT	
0	884 (88%)
1	107 (11%)
2	9 (0.9%)
WHERE_n_of_non_aggr_func__DAY	
0	949 (95%)
1	48 (4.8%)
2	2 (0.2%)
4	1 (0.1%)
WHERE_n_of_non_aggr_func__DOW	
0	966 (97%)
1	31 (3.1%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func__LOG	
0	938 (94%)
1	57 (5.7%)
2	4 (0.4%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func__ABS	
0	954 (95%)
1	45 (4.5%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__ROUND	
0	965 (97%)
1	34 (3.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__YEAR	
0	963 (96%)
1	35 (3.5%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__TRUNC	
0	962 (96%)
1	38 (3.8%)
WHERE_n_of_non_aggr_func__FLOOR	
0	961 (96%)
1	38 (3.8%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	641 (64%)
1	236 (24%)
2	91 (9.1%)
3	23 (2.3%)
4	5 (0.5%)
5	3 (0.3%)
6	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 12.00] 3.00 / 2.88

Query parameter or value	Descriptive statistics
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.49 / 2.61
HAVING_n_of_main_predicates__non_scalar_subquery	
0	427 (43%)
1	198 (20%)
2	198 (20%)
3	177 (18%)
HAVING_n_of_main_predicates__scalar_subquery	
0	410 (41%)
1	195 (20%)
2	188 (19%)
3	207 (21%)
HAVING_n_of_subqueries__non_scalar_subquery	
0	427 (43%)
1	198 (20%)
2	198 (20%)
3	177 (18%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 23.0] 2.5 / 3.7
HAVING_n_of_processed_rows_by_subquery	[0, 16, 865,056,514, 271,321,978,130] 21,837,904,087 / 36,686,130,420
ORDER_BY_n_of_columns	
0	292 (29%)
1	338 (34%)
10	1 (0.1%)
2	235 (24%)
3	102 (10%)
4	13 (1.3%)
5	10 (1.0%)
6	5 (0.5%)
7	4 (0.4%)
limit	[6, 501, 998] 495 / 294
offset	[0, 16, 1,000] 251 / 323