

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

## Scenario: 1-3jp\_ujpp\_\_g\_\_h0ss\_h0nss

## Scale factor: 0.01 GB

## The 1000-query pack was generated on: 2023-11-09

## 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, 103] 10 / 12
SELECT_n_of_non_aggr_func__LOWER	
0	712 (71%)
1	224 (22%)
2	42 (4.2%)
3	14 (1.4%)
4	2 (0.2%)
5	4 (0.4%)
7	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__LTRIM	
0	691 (69%)
1	231 (23%)
2	56 (5.6%)
3	11 (1.1%)

Query parameter or value	Descriptive statistics
4	6 (0.6%)
5	3 (0.3%)
6	2 (0.2%)
SELECT_n_of_non_aggr_func__RTRIM	
0	686 (69%)
1	235 (24%)
2	63 (6.3%)
3	12 (1.2%)
4	2 (0.2%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__SQRT	
0	856 (86%)
1	117 (12%)
2	16 (1.6%)
3	8 (0.8%)
4	3 (0.3%)
SELECT_n_of_non_aggr_func__UPPER	
0	706 (71%)
1	215 (22%)
2	59 (5.9%)
3	13 (1.3%)
4	4 (0.4%)
5	3 (0.3%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	668 (67%)
1	263 (26%)
2	54 (5.4%)
3	13 (1.3%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__FLOOR	
0	936 (94%)
1	57 (5.7%)
2	6 (0.6%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__LOG	
0	914 (91%)
1	69 (6.9%)
2	11 (1.1%)
3	6 (0.6%)
SELECT_n_of_non_aggr_func__DAY	
0	910 (91%)
1	76 (7.6%)
2	10 (1.0%)
3	3 (0.3%)
5	1 (0.1%)

Query parameter or value	Descriptive statistics
SELECT__n_of__non_aggr_func__MONTH	
0	830 (83%)
1	136 (14%)
2	22 (2.2%)
3	9 (0.9%)
4	2 (0.2%)
5	1 (0.1%)
SELECT__n_of__non_aggr_func__YEAR	
0	866 (87%)
1	107 (11%)
2	21 (2.1%)
3	5 (0.5%)
4	1 (0.1%)
SELECT__n_of__non_aggr_func__DOW	
0	885 (89%)
1	94 (9.4%)
2	18 (1.8%)
3	2 (0.2%)
5	1 (0.1%)
SELECT__n_of__non_aggr_func__ABS	
0	944 (94%)
1	51 (5.1%)
2	5 (0.5%)
SELECT__n_of__non_aggr_func__ROUND	
0	968 (97%)
1	29 (2.9%)
2	2 (0.2%)
3	1 (0.1%)
SELECT__n_of__non_aggr_func__TRUNC	
0	946 (95%)
1	50 (5.0%)
2	4 (0.4%)
SELECT__n_of__all__non_aggr_func	[0.0, 2.0, 47.0] 3.2 / 4.2
SELECT__n_of__aggr_func__AVG	
0	939 (94%)
1	43 (4.3%)
2	15 (1.5%)
3	2 (0.2%)
4	1 (0.1%)
SELECT__n_of__aggr_func__COUNT	[0.00, 0.00, 9.00] 0.66 / 1.51
SELECT__n_of__aggr_func__MAX	[0.00, 0.00, 10.00] 0.62 / 1.50
SELECT__n_of__aggr_func__MIN	[0.00, 0.00, 9.00] 0.58 / 1.33
SELECT__n_of__aggr_func__COUNT_DISTINCT	[0.00, 0.00, 10.00] 0.69 / 1.46
SELECT__n_of__aggr_func__SUM	
0	928 (93%)
1	56 (5.6%)

Query parameter or value	Descriptive statistics
2	10 (1.0%)
3	6 (0.6%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 10.00] 2.72 / 2.27
FROM_n_of_join_paths	
1	307 (31%)
2	347 (35%)
3	346 (35%)
FROM_n_of_super_joins__RIGHT	
0	776 (78%)
1	198 (20%)
2	26 (2.6%)
FROM_n_of_super_joins__LEFT	
0	757 (76%)
1	224 (22%)
2	19 (1.9%)
FROM_n_of_super_joins__FULL	
0	777 (78%)
1	193 (19%)
2	30 (3.0%)
FROM_n_of_joins__INNER	
0	210 (21%)
1	250 (25%)
2	209 (21%)
3	180 (18%)
4	87 (8.7%)
5	44 (4.4%)
6	12 (1.2%)
7	8 (0.8%)
FROM_n_of_joins__RIGHT	
0	255 (26%)
1	274 (27%)
2	197 (20%)
3	140 (14%)
4	71 (7.1%)
5	36 (3.6%)
6	20 (2.0%)
7	5 (0.5%)
8	2 (0.2%)
FROM_n_of_processed_rows	[5, 68,318, 230,115] 64,703 / 59,712
WHERE_n_of_predicates	[0.00, 3.00, 9.00] 2.82 / 2.15
WHERE_n_of_attrihs_of_type__character_varying	
0	407 (41%)
1	332 (33%)
2	179 (18%)
3	64 (6.4%)
4	15 (1.5%)

Query parameter or value	Descriptive statistics
5	3 (0.3%)
WHERE_n_of_attribs_of_type__integer	
0	445 (45%)
1	338 (34%)
2	146 (15%)
3	46 (4.6%)
4	23 (2.3%)
5	2 (0.2%)
WHERE_n_of_attribs_of_type__character	
0	731 (73%)
1	209 (21%)
2	49 (4.9%)
3	10 (1.0%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type__numeric	
0	651 (65%)
1	265 (27%)
2	66 (6.6%)
3	14 (1.4%)
4	4 (0.4%)
WHERE_n_of_attribs_of_type__date	
0	830 (83%)
1	146 (15%)
2	21 (2.1%)
3	3 (0.3%)
WHERE_n_of_pkey_attribs	
0	618 (62%)
1	279 (28%)
2	86 (8.6%)
3	15 (1.5%)
4	2 (0.2%)
WHERE_n_of_connect_OR	
0	446 (45%)
1	193 (19%)
2	140 (14%)
3	97 (9.7%)
4	56 (5.6%)
5	43 (4.3%)
6	20 (2.0%)
7	4 (0.4%)
8	1 (0.1%)
WHERE_n_of_operators__between	
0	608 (61%)
1	273 (27%)
2	99 (9.9%)
3	18 (1.8%)

Query parameter or value	Descriptive statistics
4	2 (0.2%)
WHERE_n_of_operators___greater_or_less	
0	419 (42%)
1	298 (30%)
2	185 (19%)
3	62 (6.2%)
4	24 (2.4%)
5	9 (0.9%)
6	3 (0.3%)
WHERE_n_of_operators___like	
0	821 (82%)
1	157 (16%)
2	19 (1.9%)
3	2 (0.2%)
4	1 (0.1%)
WHERE_n_of_operators___in	
0	608 (61%)
1	276 (28%)
2	98 (9.8%)
3	16 (1.6%)
4	2 (0.2%)
WHERE_n_of_non_aggr_func___LOG	
0	936 (94%)
1	63 (6.3%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___SQRT	
0	910 (91%)
1	81 (8.1%)
2	8 (0.8%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func___DOW	
0	973 (97%)
1	26 (2.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___TRUNC	
0	968 (97%)
1	32 (3.2%)
WHERE_n_of_non_aggr_func___YEAR	
0	970 (97%)
1	30 (3.0%)
WHERE_n_of_non_aggr_func___ROUND	
0	973 (97%)
1	26 (2.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___MONTH	
0	970 (97%)

Query parameter or value	Descriptive statistics
1	30 (3.0%)
WHERE_n_of_non_aggr_func__ABS	
0	970 (97%)
1	30 (3.0%)
WHERE_n_of_non_aggr_func__DAY	
0	974 (97%)
1	25 (2.5%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__FLOOR	
0	975 (98%)
1	23 (2.3%)
2	2 (0.2%)
WHERE_n_of_all_non_aggr_func	
0	694 (69%)
1	237 (24%)
2	53 (5.3%)
3	12 (1.2%)
4	2 (0.2%)
5	2 (0.2%)
GROUP_BY_n_of_columns	[0.00, 2.00, 10.00] 2.38 / 2.17
HAVING_n_of_main_predicates	
0	455 (46%)
1	176 (18%)
2	200 (20%)
3	169 (17%)
ORDER_BY_n_of_columns	
0	315 (32%)
1	392 (39%)
2	211 (21%)
3	69 (6.9%)
4	4 (0.4%)
5	6 (0.6%)
6	2 (0.2%)
8	1 (0.1%)
limit	[1, 483, 999] 492 / 289
offset	[0, 0, 993] 235 / 314