

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

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

## Scale factor: 10 GB

## The 1000-query pack was generated on: 2023-11-04

## 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	[1.0, 4.0, 40.0] 5.6 / 5.0
SELECT_n_of_non_aggr_func__ABS	
0	955 (96%)
1	44 (4.4%)
2	1 (0.1%)
SELECT_n_of_non_aggr_func__LOWER	
0	784 (78%)
1	187 (19%)
2	24 (2.4%)
3	5 (0.5%)
SELECT_n_of_non_aggr_func__LTRIM	
0	780 (78%)
1	194 (19%)
2	23 (2.3%)
3	1 (0.1%)

Query parameter or value	Descriptive statistics
4	2 (0.2%)
SELECT_n_of_non_aggr_func__LOG	
0	937 (94%)
1	59 (5.9%)
2	4 (0.4%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	791 (79%)
1	184 (18%)
2	25 (2.5%)
SELECT_n_of_non_aggr_func__YEAR	
0	913 (91%)
1	74 (7.4%)
2	12 (1.2%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__UPPER	
0	778 (78%)
1	193 (19%)
2	25 (2.5%)
3	4 (0.4%)
SELECT_n_of_non_aggr_func__RTRIM	
0	777 (78%)
1	197 (20%)
2	23 (2.3%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func__DOW	
0	922 (92%)
1	71 (7.1%)
2	7 (0.7%)
SELECT_n_of_non_aggr_func__MONTH	
0	893 (89%)
1	92 (9.2%)
2	10 (1.0%)
3	5 (0.5%)
SELECT_n_of_non_aggr_func__DAY	
0	933 (93%)
1	64 (6.4%)
2	2 (0.2%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__SQRT	
0	909 (91%)
1	83 (8.3%)
2	6 (0.6%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__FLOOR	
0	967 (97%)

Query parameter or value	Descriptive statistics
1	32 (3.2%)
2	1 (0.1%)
SELECT__n_of__non__aggr__func__ROUND	
0	972 (97%)
1	26 (2.6%)
2	2 (0.2%)
SELECT__n_of__non__aggr__func__TRUNC	
0	963 (96%)
1	36 (3.6%)
2	1 (0.1%)
SELECT__n_of__all__non__aggr__func	[0.00, 1.50, 19.00] 1.94 / 2.12
SELECT__n_of__aggr__func__MIN	
0	756 (76%)
1	117 (12%)
2	61 (6.1%)
3	39 (3.9%)
4	18 (1.8%)
5	5 (0.5%)
6	2 (0.2%)
7	2 (0.2%)
SELECT__n_of__aggr__func__MAX	
0	769 (77%)
1	110 (11%)
2	57 (5.7%)
3	35 (3.5%)
4	19 (1.9%)
5	7 (0.7%)
6	1 (0.1%)
7	1 (0.1%)
8	1 (0.1%)
SELECT__n_of__aggr__func__COUNT_DISTINCT	
0	758 (76%)
1	128 (13%)
2	56 (5.6%)
3	32 (3.2%)
4	15 (1.5%)
5	7 (0.7%)
6	2 (0.2%)
7	2 (0.2%)
SELECT__n_of__aggr__func__COUNT	
0	759 (76%)
1	103 (10%)
2	82 (8.2%)
3	32 (3.2%)
4	12 (1.2%)
5	11 (1.1%)

Query parameter or value	Descriptive statistics
6	1 (0.1%)
SELECT_n_of_aggr_func__AVG	
0	942 (94%)
1	52 (5.2%)
2	5 (0.5%)
3	1 (0.1%)
SELECT_n_of_aggr_func__SUM	
0	961 (96%)
1	32 (3.2%)
2	6 (0.6%)
3	1 (0.1%)
SELECT_n_of_all_aggr_func	
0	161 (16%)
1	272 (27%)
2	251 (25%)
3	162 (16%)
4	82 (8.2%)
5	50 (5.0%)
6	13 (1.3%)
7	7 (0.7%)
8	2 (0.2%)
FROM_n_of_join_paths	
1	1,000 (100%)
FROM_n_of_joins__RIGHT	
0	509 (51%)
1	328 (33%)
2	131 (13%)
3	29 (2.9%)
4	3 (0.3%)
FROM_n_of_joins__INNER	
0	482 (48%)
1	324 (32%)
2	152 (15%)
3	37 (3.7%)
4	5 (0.5%)
FROM_n_of_processed_rows	[5, 15,002,580, 76,480,513] 30,053,401 / 31,728,807
WHERE_n_of_predicates	
0	212 (21%)
1	244 (24%)
2	223 (22%)
3	176 (18%)
4	93 (9.3%)
5	43 (4.3%)
6	9 (0.9%)
WHERE_n_of_attris_of_type__character_varying	
0	541 (54%)

Query parameter or value	Descriptive statistics
1	341 (34%)
2	99 (9.9%)
3	16 (1.6%)
4	3 (0.3%)
WHERE_n_of_attribs_of_type__character	
0	784 (78%)
1	190 (19%)
2	22 (2.2%)
3	3 (0.3%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type__integer	
0	545 (55%)
1	348 (35%)
2	93 (9.3%)
3	12 (1.2%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type__numeric	
0	752 (75%)
1	196 (20%)
2	45 (4.5%)
3	7 (0.7%)
WHERE_n_of_attribs_of_type__date	
0	881 (88%)
1	110 (11%)
2	9 (0.9%)
WHERE_n_of_pkey_attribs	
0	702 (70%)
1	259 (26%)
2	35 (3.5%)
3	4 (0.4%)
WHERE_n_of_connect_OR	
0	577 (58%)
1	242 (24%)
2	119 (12%)
3	46 (4.6%)
4	14 (1.4%)
5	2 (0.2%)
WHERE_n_of_operators__greater_or_less	
0	512 (51%)
1	306 (31%)
2	138 (14%)
3	39 (3.9%)
4	5 (0.5%)
WHERE_n_of_operators__between	
0	717 (72%)
1	238 (24%)

Query parameter or value	Descriptive statistics
2	39 (3.9%)
3	6 (0.6%)
WHERE_n_of_operators_in	
0	707 (71%)
1	250 (25%)
2	41 (4.1%)
3	2 (0.2%)
WHERE_n_of_operators_like	
0	874 (87%)
1	119 (12%)
2	6 (0.6%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func_LOG	
0	951 (95%)
1	49 (4.9%)
WHERE_n_of_non_aggr_func_SQRT	
0	941 (94%)
1	58 (5.8%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_YEAR	
0	987 (99%)
1	13 (1.3%)
WHERE_n_of_non_aggr_func_FLOOR	
0	980 (98%)
1	20 (2.0%)
WHERE_n_of_non_aggr_func_TRUNC	
0	981 (98%)
1	18 (1.8%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_ROUND	
0	976 (98%)
1	24 (2.4%)
WHERE_n_of_non_aggr_func_DOW	
0	981 (98%)
1	19 (1.9%)
WHERE_n_of_non_aggr_func_ABS	
0	984 (98%)
1	14 (1.4%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func_DAY	
0	979 (98%)
1	21 (2.1%)
WHERE_n_of_non_aggr_func_MONTH	
0	978 (98%)
1	20 (2.0%)
2	2 (0.2%)

Query parameter or value	Descriptive statistics
WHERE_n_of_all_non_aggr_func	
0	775 (78%)
1	185 (19%)
2	37 (3.7%)
3	3 (0.3%)
GROUP_BY_n_of_columns	
0	222 (22%)
1	270 (27%)
2	238 (24%)
3	157 (16%)
4	76 (7.6%)
5	31 (3.1%)
6	6 (0.6%)
HAVING_n_of_main_predicates	
0	434 (43%)
1	182 (18%)
2	194 (19%)
3	190 (19%)
ORDER_BY_n_of_columns	
0	328 (33%)
1	446 (45%)
2	196 (20%)
3	22 (2.2%)
4	8 (0.8%)
limit	[2, 492, 1,000] 496 / 290
offset	[0, 0, 1,000] 254 / 325