

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

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

## Scale factor: 10 GB

## The 1000-query pack was generated on: 2023-11-10

## 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, 99] 10 / 13
SELECT_n_of_non_aggr_func__DAY	
0	922 (92%)
1	63 (6.3%)
2	10 (1.0%)
3	4 (0.4%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__MONTH	
0	850 (85%)
1	115 (12%)
2	25 (2.5%)
3	5 (0.5%)
4	3 (0.3%)
5	1 (0.1%)
7	1 (0.1%)

Query parameter or value	Descriptive statistics
SELECT__n_of__non_aggr_func__LOWER	
0	701 (70%)
1	227 (23%)
2	56 (5.6%)
3	6 (0.6%)
4	6 (0.6%)
5	2 (0.2%)
6	2 (0.2%)
SELECT__n_of__non_aggr_func__RTRIM	
0	697 (70%)
1	226 (23%)
2	59 (5.9%)
3	12 (1.2%)
4	4 (0.4%)
5	1 (0.1%)
6	1 (0.1%)
SELECT__n_of__non_aggr_func__TRUNC	
0	945 (95%)
1	50 (5.0%)
2	3 (0.3%)
3	2 (0.2%)
SELECT__n_of__non_aggr_func__UPPER	
0	669 (67%)
1	256 (26%)
2	59 (5.9%)
3	11 (1.1%)
4	4 (0.4%)
5	1 (0.1%)
SELECT__n_of__non_aggr_func__SUBSTR	
0	702 (70%)
1	232 (23%)
2	47 (4.7%)
3	11 (1.1%)
4	6 (0.6%)
5	1 (0.1%)
6	1 (0.1%)
SELECT__n_of__non_aggr_func__LTRIM	
0	694 (69%)
1	238 (24%)
2	51 (5.1%)
3	11 (1.1%)
4	5 (0.5%)
5	1 (0.1%)
SELECT__n_of__non_aggr_func__LOG	
0	897 (90%)
1	88 (8.8%)

Query parameter or value	Descriptive statistics
2	9 (0.9%)
3	5 (0.5%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__FLOOR	
0	938 (94%)
1	57 (5.7%)
2	4 (0.4%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__SQRT	
0	858 (86%)
1	102 (10%)
2	28 (2.8%)
3	9 (0.9%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__ROUND	
0	953 (95%)
1	37 (3.7%)
2	7 (0.7%)
3	2 (0.2%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__YEAR	
0	874 (87%)
1	96 (9.6%)
2	25 (2.5%)
3	4 (0.4%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__DOW	
0	879 (88%)
1	96 (9.6%)
2	19 (1.9%)
3	4 (0.4%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__ABS	
0	937 (94%)
1	58 (5.8%)
2	5 (0.5%)
SELECT_n_of_all_non_aggr_func	[0.0, 2.0, 34.0] 3.2 / 4.2
SELECT_n_of_aggr_func__COUNT_DISTINCT	
0	760 (76%)
1	79 (7.9%)
2	61 (6.1%)
3	41 (4.1%)
4	25 (2.5%)
5	14 (1.4%)
6	11 (1.1%)

Query parameter or value	Descriptive statistics
7	5 (0.5%)
8	4 (0.4%)
SELECT__n_of_aggr_func__COUNT	[0.00, 0.00, 11.00] 0.63 / 1.44
SELECT__n_of_aggr_func__MAX	[0.00, 0.00, 9.00] 0.60 / 1.36
SELECT__n_of_aggr_func__MIN	[0.00, 0.00, 10.00] 0.59 / 1.49
SELECT__n_of_aggr_func__AVG	
0	942 (94%)
1	47 (4.7%)
2	8 (0.8%)
3	3 (0.3%)
SELECT__n_of_aggr_func__SUM	
0	942 (94%)
1	43 (4.3%)
2	12 (1.2%)
3	3 (0.3%)
SELECT__n_of_all_aggr_func	[0.00, 2.00, 11.00] 2.59 / 2.27
FROM__n_of_join_paths	
1	335 (34%)
2	319 (32%)
3	346 (35%)
FROM__n_of_super_joins__FULL	
0	770 (77%)
1	207 (21%)
2	23 (2.3%)
FROM__n_of_super_joins__RIGHT	
0	769 (77%)
1	206 (21%)
2	25 (2.5%)
FROM__n_of_super_joins__LEFT	
0	762 (76%)
1	219 (22%)
2	19 (1.9%)
FROM__n_of_joins__INNER	
0	250 (25%)
1	241 (24%)
2	221 (22%)
3	150 (15%)
4	76 (7.6%)
5	33 (3.3%)
6	25 (2.5%)
7	4 (0.4%)
FROM__n_of_joins__RIGHT	
0	239 (24%)
1	264 (26%)
2	209 (21%)
3	141 (14%)

Query parameter or value	Descriptive statistics
4	88 (8.8%)
5	39 (3.9%)
6	16 (1.6%)
7	3 (0.3%)
8	1 (0.1%)
FROM_n_of_processed_rows	[5, 68,079,151, 229,441,539] 64,865,656 / 59,769,286
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.73 / 2.18
WHERE_n_of_attribs_of_type___character_varying	
0	457 (46%)
1	273 (27%)
2	161 (16%)
3	82 (8.2%)
4	20 (2.0%)
5	7 (0.7%)
WHERE_n_of_attribs_of_type___integer	
0	482 (48%)
1	321 (32%)
2	145 (15%)
3	45 (4.5%)
4	5 (0.5%)
5	2 (0.2%)
WHERE_n_of_attribs_of_type___numeric	
0	664 (66%)
1	248 (25%)
2	68 (6.8%)
3	17 (1.7%)
4	3 (0.3%)
WHERE_n_of_attribs_of_type___character	
0	719 (72%)
1	222 (22%)
2	51 (5.1%)
3	7 (0.7%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type___date	
0	830 (83%)
1	144 (14%)
2	25 (2.5%)
4	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	644 (64%)
1	272 (27%)
2	73 (7.3%)
3	11 (1.1%)
WHERE_n_of_connect_OR	[0.00, 1.00, 9.00] 1.33 / 1.66
WHERE_n_of_operators___between	
0	603 (60%)

Query parameter or value	Descriptive statistics
1	283 (28%)
2	94 (9.4%)
3	18 (1.8%)
4	2 (0.2%)
WHERE_n_of_operators___greater_or_less	
0	421 (42%)
1	309 (31%)
2	165 (17%)
3	67 (6.7%)
4	28 (2.8%)
5	10 (1.0%)
WHERE_n_of_operators___like	
0	823 (82%)
1	161 (16%)
2	14 (1.4%)
3	2 (0.2%)
WHERE_n_of_operators___in	
0	662 (66%)
1	232 (23%)
2	79 (7.9%)
3	23 (2.3%)
4	2 (0.2%)
5	2 (0.2%)
WHERE_n_of_non_aggr_func___MONTH	
0	977 (98%)
1	23 (2.3%)
WHERE_n_of_non_aggr_func___TRUNC	
0	972 (97%)
1	28 (2.8%)
WHERE_n_of_non_aggr_func___LOG	
0	942 (94%)
1	56 (5.6%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___SQRT	
0	925 (93%)
1	68 (6.8%)
2	6 (0.6%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func___ROUND	
0	962 (96%)
1	37 (3.7%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___DAY	
0	966 (97%)
1	34 (3.4%)
WHERE_n_of_non_aggr_func___FLOOR	

Query parameter or value	Descriptive statistics
0	967 (97%)
1	33 (3.3%)
WHERE_n_of_non_aggr_func__ABS	
0	986 (99%)
1	13 (1.3%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__YEAR	
0	977 (98%)
1	22 (2.2%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DOW	
0	967 (97%)
1	33 (3.3%)
WHERE_n_of_all_non_aggr_func	
0	714 (71%)
1	212 (21%)
2	63 (6.3%)
3	10 (1.0%)
4	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.36 / 2.17
HAVING_n_of_main_predicates	
0	430 (43%)
1	195 (20%)
2	188 (19%)
3	187 (19%)
ORDER_BY_n_of_columns	
0	307 (31%)
1	372 (37%)
2	239 (24%)
3	67 (6.7%)
4	7 (0.7%)
5	7 (0.7%)
6	1 (0.1%)
limit	[1, 508, 998] 502 / 289
offset	[0, 17, 997] 259 / 324