

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

## Scenario: 1-6jp\_ujpp\_\_g\_\_h0ss\_h0nss

## Scale factor: 1000 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, 8, 169] 15 / 22
SELECT_n_of_non_aggr_func__SQRT	
0	796 (80%)
1	144 (14%)
2	35 (3.5%)
3	12 (1.2%)
4	6 (0.6%)
5	4 (0.4%)
6	1 (0.1%)
7	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__UPPER	
0	636 (64%)
1	252 (25%)
2	70 (7.0%)

Query parameter or value	Descriptive statistics
3	20 (2.0%)
4	9 (0.9%)
5	9 (0.9%)
6	2 (0.2%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func__LOG	
0	870 (87%)
1	104 (10%)
2	14 (1.4%)
3	9 (0.9%)
4	1 (0.1%)
6	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__DOW	
0	867 (87%)
1	98 (9.8%)
2	23 (2.3%)
3	4 (0.4%)
4	4 (0.4%)
5	2 (0.2%)
6	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__LOWER	
0	603 (60%)
1	267 (27%)
2	87 (8.7%)
3	24 (2.4%)
4	9 (0.9%)
5	6 (0.6%)
6	3 (0.3%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__LTRIM	
0	632 (63%)
1	260 (26%)
10	1 (0.1%)
2	72 (7.2%)
3	22 (2.2%)
4	6 (0.6%)
5	4 (0.4%)
6	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__MONTH	
0	826 (83%)
1	115 (12%)
2	32 (3.2%)
3	14 (1.4%)

Query parameter or value	Descriptive statistics
4	8 (0.8%)
5	2 (0.2%)
6	2 (0.2%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__DAY	
0	887 (89%)
1	84 (8.4%)
2	17 (1.7%)
3	9 (0.9%)
4	1 (0.1%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func__RTRIM	[0.00, 0.00, 10.00] 0.57 / 1.03
SELECT_n_of_non_aggr_func__SUBSTR	
0	639 (64%)
1	224 (22%)
2	97 (9.7%)
3	20 (2.0%)
4	9 (0.9%)
6	6 (0.6%)
7	3 (0.3%)
8	1 (0.1%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func__YEAR	
0	834 (83%)
1	119 (12%)
2	27 (2.7%)
3	12 (1.2%)
4	6 (0.6%)
5	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__ABS	
0	920 (92%)
1	67 (6.7%)
2	8 (0.8%)
3	4 (0.4%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__ROUND	
0	921 (92%)
1	64 (6.4%)
2	11 (1.1%)
3	1 (0.1%)
4	3 (0.3%)
SELECT_n_of_non_aggr_func__TRUNC	
0	923 (92%)
1	67 (6.7%)
2	7 (0.7%)

Query parameter or value	Descriptive statistics
3	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__FLOOR	
0	933 (93%)
1	56 (5.6%)
2	10 (1.0%)
3	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 70.0] 4.6 / 7.1
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 13.00] 0.85 / 1.89
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 10.00] 0.73 / 1.74
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 13.00] 0.93 / 2.04
SELECT_n_of_aggr_func__AVG	
0	946 (95%)
1	32 (3.2%)
2	16 (1.6%)
3	5 (0.5%)
4	1 (0.1%)
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 12.00] 0.84 / 1.93
SELECT_n_of_aggr_func__SUM	
0	934 (93%)
1	45 (4.5%)
2	14 (1.4%)
3	5 (0.5%)
4	2 (0.2%)
SELECT_n_of_all_aggr_func	[0.00, 3.00, 14.00] 3.54 / 3.08
FROM_n_of_join_paths	
1	179 (18%)
2	173 (17%)
3	164 (16%)
4	145 (15%)
5	162 (16%)
6	177 (18%)
FROM_n_of_super_joins__FULL	
0	544 (54%)
1	312 (31%)
2	114 (11%)
3	24 (2.4%)
4	6 (0.6%)
FROM_n_of_super_joins__LEFT	
0	562 (56%)
1	288 (29%)
2	118 (12%)
3	28 (2.8%)
4	4 (0.4%)
FROM_n_of_super_joins__RIGHT	
0	542 (54%)

Query parameter or value	Descriptive statistics
1	318 (32%)
2	114 (11%)
3	23 (2.3%)
4	3 (0.3%)
FROM_n_of_joins__INNER	[0.00, 3.00, 14.00] 3.27 / 2.54
FROM_n_of_joins__RIGHT	[0.00, 3.00, 15.00] 3.30 / 2.52
FROM_n_of_processed_rows	[5, 9,244,989,754, 44,059,938,309] 11,624,405,085 / 9,755,323,523
WHERE_n_of_predicates	[0.00, 3.00, 12.00] 3.84 / 2.93
WHERE_n_of_attribs_of_type__character_varying	
0	371 (37%)
1	301 (30%)
2	184 (18%)
3	87 (8.7%)
4	38 (3.8%)
5	13 (1.3%)
6	6 (0.6%)
WHERE_n_of_attribs_of_type__character	
0	638 (64%)
1	248 (25%)
2	85 (8.5%)
3	25 (2.5%)
4	4 (0.4%)
WHERE_n_of_attribs_of_type__date	
0	762 (76%)
1	192 (19%)
2	39 (3.9%)
3	4 (0.4%)
4	3 (0.3%)
WHERE_n_of_attribs_of_type__numeric	
0	543 (54%)
1	291 (29%)
2	120 (12%)
3	29 (2.9%)
4	16 (1.6%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type__integer	
0	372 (37%)
1	304 (30%)
2	176 (18%)
3	94 (9.4%)
4	40 (4.0%)
5	11 (1.1%)
6	3 (0.3%)
WHERE_n_of_pkey_attribs	
0	535 (54%)

Query parameter or value	Descriptive statistics
1	323 (32%)
2	108 (11%)
3	29 (2.9%)
4	3 (0.3%)
5	2 (0.2%)
WHERE_n_of_connect_OR	[0.00, 1.00, 10.00] 2.21 / 2.44
WHERE_n_of_operators___greater_or_less	
0	343 (34%)
1	258 (26%)
2	194 (19%)
3	102 (10%)
4	63 (6.3%)
5	31 (3.1%)
6	6 (0.6%)
7	3 (0.3%)
WHERE_n_of_operators___in	
0	535 (54%)
1	275 (28%)
2	117 (12%)
3	61 (6.1%)
4	9 (0.9%)
5	3 (0.3%)
WHERE_n_of_operators___like	
0	797 (80%)
1	172 (17%)
2	25 (2.5%)
3	6 (0.6%)
WHERE_n_of_operators___between	
0	534 (53%)
1	286 (29%)
2	130 (13%)
3	39 (3.9%)
4	9 (0.9%)
5	1 (0.1%)
6	1 (0.1%)
WHERE_n_of_non_aggr_func___ROUND	
0	948 (95%)
1	50 (5.0%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___YEAR	
0	956 (96%)
1	43 (4.3%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___ABS	
0	969 (97%)
1	28 (2.8%)

Query parameter or value	Descriptive statistics
2	3 (0.3%)
WHERE_n_of_non_aggr_func__DAY	
0	963 (96%)
1	37 (3.7%)
WHERE_n_of_non_aggr_func__LOG	
0	922 (92%)
1	75 (7.5%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func__FLOOR	
0	949 (95%)
1	50 (5.0%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__SQRT	
0	876 (88%)
1	114 (11%)
2	8 (0.8%)
3	2 (0.2%)
WHERE_n_of_non_aggr_func__DOW	
0	957 (96%)
1	42 (4.2%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__TRUNC	
0	956 (96%)
1	40 (4.0%)
2	4 (0.4%)
WHERE_n_of_non_aggr_func__MONTH	
0	960 (96%)
1	40 (4.0%)
WHERE_n_of_all_non_aggr_func	
0	620 (62%)
1	245 (25%)
2	89 (8.9%)
3	37 (3.7%)
4	8 (0.8%)
5	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 13.00] 3.11 / 2.93
HAVING_n_of_main_predicates	
0	412 (41%)
1	180 (18%)
2	201 (20%)
3	207 (21%)
ORDER_BY_n_of_columns	
0	292 (29%)
1	345 (35%)
2	236 (24%)
3	96 (9.6%)

Query parameter or value	Descriptive statistics
4	24 (2.4%)
5	4 (0.4%)
6	1 (0.1%)
7	2 (0.2%)
limit	[1, 481, 1,000] 492 / 289
offset	[0, 4, 994] 253 / 325