

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

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

## Scale factor: 200 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, 91] 10 / 12
SELECT_n_of_non_aggr_func__SQRT	
0	855 (86%)
1	124 (12%)
2	11 (1.1%)
3	6 (0.6%)
4	3 (0.3%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__TRUNC	
0	929 (93%)
1	63 (6.3%)
2	7 (0.7%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__UPPER	
0	724 (72%)

Query parameter or value	Descriptive statistics
1	209 (21%)
2	50 (5.0%)
3	9 (0.9%)
4	6 (0.6%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__LOWER	
0	683 (68%)
1	234 (23%)
2	64 (6.4%)
3	12 (1.2%)
4	3 (0.3%)
5	4 (0.4%)
SELECT_n_of_non_aggr_func__LTRIM	
0	700 (70%)
1	228 (23%)
2	52 (5.2%)
3	14 (1.4%)
4	6 (0.6%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	686 (69%)
1	230 (23%)
2	62 (6.2%)
3	16 (1.6%)
4	5 (0.5%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__LOG	
0	884 (88%)
1	101 (10%)
2	11 (1.1%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__RTRIM	
0	681 (68%)
1	233 (23%)
2	61 (6.1%)
3	14 (1.4%)
4	8 (0.8%)
5	3 (0.3%)
SELECT_n_of_non_aggr_func__MONTH	
0	859 (86%)
1	97 (9.7%)
2	28 (2.8%)
3	10 (1.0%)
4	3 (0.3%)
5	1 (0.1%)

Query parameter or value	Descriptive statistics
6	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__YEAR	
0	865 (87%)
1	102 (10%)
2	25 (2.5%)
3	6 (0.6%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__DAY	
0	918 (92%)
1	67 (6.7%)
2	12 (1.2%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func__ABS	
0	941 (94%)
1	56 (5.6%)
2	2 (0.2%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__DOW	
0	882 (88%)
1	97 (9.7%)
2	11 (1.1%)
3	7 (0.7%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__ROUND	
0	946 (95%)
1	46 (4.6%)
2	8 (0.8%)
SELECT_n_of_non_aggr_func__FLOOR	
0	947 (95%)
1	52 (5.2%)
2	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.0, 2.0, 36.0] 3.3 / 4.4
SELECT_n_of_aggr_func__AVG	
0	946 (95%)
1	39 (3.9%)
2	11 (1.1%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 10.00] 0.66 / 1.46
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 9.00] 0.65 / 1.46
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 12.00] 0.63 / 1.44
SELECT_n_of_aggr_func__MIN	
0	769 (77%)
1	77 (7.7%)

Query parameter or value	Descriptive statistics
2	59 (5.9%)
3	27 (2.7%)
4	25 (2.5%)
5	19 (1.9%)
6	13 (1.3%)
7	8 (0.8%)
8	3 (0.3%)
SELECT_n_of_aggr_func__SUM	
0	940 (94%)
1	44 (4.4%)
2	11 (1.1%)
3	4 (0.4%)
4	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 12.00] 2.73 / 2.28
FROM_n_of_join_paths	
1	328 (33%)
2	336 (34%)
3	336 (34%)
FROM_n_of_super_joins__FULL	
0	752 (75%)
1	228 (23%)
2	20 (2.0%)
FROM_n_of_super_joins__LEFT	
0	776 (78%)
1	205 (21%)
2	19 (1.9%)
FROM_n_of_super_joins__RIGHT	
0	769 (77%)
1	210 (21%)
2	21 (2.1%)
FROM_n_of_joins__INNER	
0	243 (24%)
1	276 (28%)
2	208 (21%)
3	124 (12%)
4	99 (9.9%)
5	36 (3.6%)
6	9 (0.9%)
7	5 (0.5%)
FROM_n_of_joins__RIGHT	
0	235 (24%)
1	283 (28%)
2	210 (21%)
3	142 (14%)
4	81 (8.1%)
5	28 (2.8%)

Query parameter or value	Descriptive statistics
6	14 (1.4%)
7	6 (0.6%)
8	1 (0.1%)
FROM_n_of_processed_rows	[5, 1,362,018,464, 4,590,055,357] 1,278,542,612 / 1,167,414,662
WHERE_n_of_predicates	[0.00, 3.00, 9.00] 2.83 / 2.21
WHERE_n_of_attribs_of_type__date	
0	833 (83%)
1	146 (15%)
2	18 (1.8%)
3	3 (0.3%)
WHERE_n_of_attribs_of_type__integer	
0	458 (46%)
1	304 (30%)
2	170 (17%)
3	46 (4.6%)
4	15 (1.5%)
5	6 (0.6%)
6	1 (0.1%)
WHERE_n_of_attribs_of_type__numeric	
0	643 (64%)
1	264 (26%)
2	76 (7.6%)
3	14 (1.4%)
4	2 (0.2%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type__character_varying	
0	414 (41%)
1	326 (33%)
2	170 (17%)
3	70 (7.0%)
4	18 (1.8%)
5	2 (0.2%)
WHERE_n_of_attribs_of_type__character	
0	731 (73%)
1	210 (21%)
2	51 (5.1%)
3	8 (0.8%)
WHERE_n_of_pkey_attribs	
0	613 (61%)
1	296 (30%)
2	80 (8.0%)
3	7 (0.7%)
4	4 (0.4%)
WHERE_n_of_connect_OR	
0	446 (45%)

Query parameter or value	Descriptive statistics
1	177 (18%)
2	138 (14%)
3	110 (11%)
4	63 (6.3%)
5	39 (3.9%)
6	22 (2.2%)
7	5 (0.5%)
WHERE_n_of_operators___greater_or_less	
0	410 (41%)
1	296 (30%)
2	165 (17%)
3	81 (8.1%)
4	33 (3.3%)
5	11 (1.1%)
6	3 (0.3%)
7	1 (0.1%)
WHERE_n_of_operators___in	
0	598 (60%)
1	288 (29%)
2	87 (8.7%)
3	25 (2.5%)
4	1 (0.1%)
5	1 (0.1%)
WHERE_n_of_operators___between	
0	628 (63%)
1	272 (27%)
2	78 (7.8%)
3	14 (1.4%)
4	8 (0.8%)
WHERE_n_of_operators___like	
0	823 (82%)
1	157 (16%)
2	20 (2.0%)
WHERE_n_of_non_aggr_func___ROUND	
0	968 (97%)
1	32 (3.2%)
WHERE_n_of_non_aggr_func___YEAR	
0	971 (97%)
1	29 (2.9%)
WHERE_n_of_non_aggr_func___TRUNC	
0	976 (98%)
1	24 (2.4%)
WHERE_n_of_non_aggr_func___ABS	
0	973 (97%)
1	26 (2.6%)
2	1 (0.1%)

Query parameter or value	Descriptive statistics
WHERE_n_of_non_aggr_func__DOW	
0	977 (98%)
1	21 (2.1%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__FLOOR	
0	975 (98%)
1	25 (2.5%)
WHERE_n_of_non_aggr_func__LOG	
0	925 (93%)
1	73 (7.3%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__SQRT	
0	908 (91%)
1	88 (8.8%)
2	4 (0.4%)
WHERE_n_of_non_aggr_func__MONTH	
0	974 (97%)
1	26 (2.6%)
WHERE_n_of_non_aggr_func__DAY	
0	975 (98%)
1	25 (2.5%)
WHERE_n_of_all_non_aggr_func	
0	695 (70%)
1	240 (24%)
2	48 (4.8%)
3	17 (1.7%)
GROUP_BY_n_of_columns	[0.00, 2.00, 10.00] 2.35 / 2.10
HAVING_n_of_main_predicates	
0	445 (45%)
1	191 (19%)
2	181 (18%)
3	183 (18%)
ORDER_BY_n_of_columns	
0	292 (29%)
1	392 (39%)
2	231 (23%)
3	60 (6.0%)
4	17 (1.7%)
5	4 (0.4%)
6	2 (0.2%)
7	2 (0.2%)
limit	[1, 490, 999] 495 / 292
offset	[0, 0, 1,000] 236 / 316