

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

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

## Scale factor: 0.01 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, 113] 10 / 12
SELECT_n_of_non_aggr_func__LOWER	
0	676 (68%)
1	249 (25%)
2	61 (6.1%)
3	12 (1.2%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	704 (70%)
1	226 (23%)
2	49 (4.9%)
3	17 (1.7%)
4	2 (0.2%)
5	1 (0.1%)
6	1 (0.1%)

Query parameter or value	Descriptive statistics
SELECT__n_of__non_aggr_func__UPPER	
0	693 (69%)
1	233 (23%)
2	55 (5.5%)
3	12 (1.2%)
4	4 (0.4%)
5	3 (0.3%)
SELECT__n_of__non_aggr_func__ABS	
0	943 (94%)
1	50 (5.0%)
2	6 (0.6%)
3	1 (0.1%)
SELECT__n_of__non_aggr_func__ROUND	
0	936 (94%)
1	59 (5.9%)
2	5 (0.5%)
SELECT__n_of__non_aggr_func__DOW	
0	887 (89%)
1	95 (9.5%)
2	11 (1.1%)
3	4 (0.4%)
4	1 (0.1%)
5	2 (0.2%)
SELECT__n_of__non_aggr_func__RTRIM	
0	724 (72%)
1	210 (21%)
2	46 (4.6%)
3	14 (1.4%)
4	5 (0.5%)
5	1 (0.1%)
SELECT__n_of__non_aggr_func__SQRT	
0	853 (85%)
1	115 (12%)
2	18 (1.8%)
3	12 (1.2%)
4	2 (0.2%)
SELECT__n_of__non_aggr_func__YEAR	
0	862 (86%)
1	106 (11%)
2	23 (2.3%)
3	6 (0.6%)
4	2 (0.2%)
5	1 (0.1%)
SELECT__n_of__non_aggr_func__LTRIM	
0	697 (70%)
1	233 (23%)

Query parameter or value	Descriptive statistics
2	48 (4.8%)
3	15 (1.5%)
4	5 (0.5%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__DAY	
0	909 (91%)
1	78 (7.8%)
2	12 (1.2%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__MONTH	
0	845 (85%)
1	123 (12%)
2	15 (1.5%)
3	11 (1.1%)
4	5 (0.5%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__LOG	
0	886 (89%)
1	101 (10%)
2	13 (1.3%)
SELECT_n_of_non_aggr_func__FLOOR	
0	948 (95%)
1	46 (4.6%)
2	5 (0.5%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__TRUNC	
0	952 (95%)
1	41 (4.1%)
2	6 (0.6%)
4	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.00, 2.00, 43.00] 3.21 / 4.17
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 9.00] 0.63 / 1.46
SELECT_n_of_aggr_func__AVG	
0	935 (94%)
1	42 (4.2%)
2	21 (2.1%)
3	2 (0.2%)
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 9.00] 0.72 / 1.53
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 9.00] 0.61 / 1.42
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 10.00] 0.61 / 1.36
SELECT_n_of_aggr_func__SUM	
0	942 (94%)
1	44 (4.4%)
2	11 (1.1%)
3	3 (0.3%)

Query parameter or value	Descriptive statistics
SELECT_n_of_all_aggr_func	[0.00, 2.00, 11.00] 2.74 / 2.29
FROM_n_of_join_paths	
1	343 (34%)
2	338 (34%)
3	319 (32%)
FROM_n_of_super_joins__RIGHT	
0	773 (77%)
1	205 (21%)
2	22 (2.2%)
FROM_n_of_super_joins__LEFT	
0	794 (79%)
1	184 (18%)
2	22 (2.2%)
FROM_n_of_super_joins__FULL	
0	774 (77%)
1	206 (21%)
2	20 (2.0%)
FROM_n_of_joins__INNER	[0.00, 1.00, 10.00] 1.74 / 1.57
FROM_n_of_joins__RIGHT	
0	252 (25%)
1	260 (26%)
2	219 (22%)
3	139 (14%)
4	67 (6.7%)
5	42 (4.2%)
6	18 (1.8%)
7	3 (0.3%)
FROM_n_of_processed_rows	[5, 68,275, 230,100] 63,383 / 58,757
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.79 / 2.19
WHERE_n_of_attribs_of_type__character_varying	
0	430 (43%)
1	317 (32%)
2	170 (17%)
3	64 (6.4%)
4	15 (1.5%)
5	3 (0.3%)
6	1 (0.1%)
WHERE_n_of_attribs_of_type__integer	
0	449 (45%)
1	337 (34%)
2	141 (14%)
3	50 (5.0%)
4	21 (2.1%)
5	1 (0.1%)
7	1 (0.1%)
WHERE_n_of_attribs_of_type__numeric	

Query parameter or value	Descriptive statistics
0	649 (65%)
1	253 (25%)
2	75 (7.5%)
3	20 (2.0%)
4	3 (0.3%)
WHERE_n_of_attribs_of_type__character	
0	727 (73%)
1	223 (22%)
2	37 (3.7%)
3	11 (1.1%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type__date	
0	846 (85%)
1	128 (13%)
2	23 (2.3%)
3	3 (0.3%)
WHERE_n_of_pkey_attribs	
0	611 (61%)
1	288 (29%)
2	83 (8.3%)
3	13 (1.3%)
4	5 (0.5%)
WHERE_n_of_connect_OR	
0	456 (46%)
1	186 (19%)
2	142 (14%)
3	93 (9.3%)
4	57 (5.7%)
5	35 (3.5%)
6	21 (2.1%)
7	8 (0.8%)
8	2 (0.2%)
WHERE_n_of_operators__between	
0	618 (62%)
1	270 (27%)
2	86 (8.6%)
3	24 (2.4%)
4	2 (0.2%)
WHERE_n_of_operators__greater_or_less	
0	416 (42%)
1	303 (30%)
2	165 (17%)
3	77 (7.7%)
4	25 (2.5%)
5	10 (1.0%)
6	4 (0.4%)

Query parameter or value	Descriptive statistics
WHERE_n_of_operators_in	
0	613 (61%)
1	279 (28%)
2	85 (8.5%)
3	18 (1.8%)
4	4 (0.4%)
5	1 (0.1%)
WHERE_n_of_operators_like	
0	842 (84%)
1	141 (14%)
2	16 (1.6%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func_TRUNC	
0	965 (97%)
1	33 (3.3%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func_SQRT	
0	916 (92%)
1	81 (8.1%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func_FLOOR	
0	960 (96%)
1	39 (3.9%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_LOG	
0	936 (94%)
1	59 (5.9%)
2	5 (0.5%)
WHERE_n_of_non_aggr_func_ABS	
0	965 (97%)
1	34 (3.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_ROUND	
0	966 (97%)
1	34 (3.4%)
WHERE_n_of_non_aggr_func_DOW	
0	978 (98%)
1	20 (2.0%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func_DAY	
0	980 (98%)
1	20 (2.0%)
WHERE_n_of_non_aggr_func_MONTH	
0	980 (98%)
1	20 (2.0%)
WHERE_n_of_non_aggr_func_YEAR	

Query parameter or value	Descriptive statistics
<i>0</i>	972 (97%)
<i>1</i>	28 (2.8%)
WHERE_n_of_all_non_aggr_func	
<i>0</i>	701 (70%)
<i>1</i>	216 (22%)
<i>2</i>	70 (7.0%)
<i>3</i>	12 (1.2%)
<i>4</i>	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.44 / 2.14
HAVING_n_of_main_predicates	
<i>0</i>	442 (44%)
<i>1</i>	174 (17%)
<i>2</i>	193 (19%)
<i>3</i>	191 (19%)
ORDER_BY_n_of_columns	
<i>0</i>	306 (31%)
<i>1</i>	350 (35%)
<i>2</i>	265 (27%)
<i>3</i>	60 (6.0%)
<i>4</i>	10 (1.0%)
<i>5</i>	7 (0.7%)
<i>6</i>	1 (0.1%)
<i>8</i>	1 (0.1%)
limit	[1, 477, 1,000] 486 / 290
offset	[0, 0, 995] 247 / 318