

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

```
## Scenario: 1-6jp_rjpp__g__h0-3ss_h0nss

## Scale factor: 2000 GB

## The 1000-query pack was generated on: 2023-11-07

## 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, 9, 201] 15 / 23
SELECT_n_of_non_aggr_func__LOWER	
0	607 (61%)
1	277 (28%)
2	74 (7.4%)
3	22 (2.2%)
4	9 (0.9%)
5	6 (0.6%)
6	3 (0.3%)
7	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__ROUND	
0	918 (92%)
1	65 (6.5%)
2	13 (1.3%)

Query parameter or value	Descriptive statistics
3	4 (0.4%)
SELECT__n_of__non__aggr__func__LOG	
0	853 (85%)
1	114 (11%)
2	22 (2.2%)
3	6 (0.6%)
4	2 (0.2%)
5	2 (0.2%)
6	1 (0.1%)
SELECT__n_of__non__aggr__func__LTRIM	[0.00, 0.00, 12.00] 0.65 / 1.10
SELECT__n_of__non__aggr__func__RTRIM	
0	593 (59%)
1	281 (28%)
2	82 (8.2%)
3	21 (2.1%)
4	14 (1.4%)
5	2 (0.2%)
6	3 (0.3%)
7	3 (0.3%)
8	1 (0.1%)
SELECT__n_of__non__aggr__func__SUBSTR	
0	614 (61%)
1	254 (25%)
2	77 (7.7%)
3	31 (3.1%)
4	11 (1.1%)
5	10 (1.0%)
6	2 (0.2%)
8	1 (0.1%)
SELECT__n_of__non__aggr__func__DOW	
0	834 (83%)
1	123 (12%)
2	25 (2.5%)
3	10 (1.0%)
4	6 (0.6%)
5	2 (0.2%)
SELECT__n_of__non__aggr__func__MONTH	[0.00, 0.00, 11.00] 0.35 / 0.91
SELECT__n_of__non__aggr__func__SQRT	
0	779 (78%)
1	150 (15%)
2	38 (3.8%)
3	8 (0.8%)
4	6 (0.6%)
5	10 (1.0%)
6	4 (0.4%)
7	4 (0.4%)

Query parameter or value	Descriptive statistics
8	1 (0.1%)
SELECT_n_of_non_aggr_func__UPPER	[0.00, 0.00, 11.00] 0.64 / 1.12
SELECT_n_of_non_aggr_func__YEAR	
0	801 (80%)
1	143 (14%)
2	34 (3.4%)
3	9 (0.9%)
4	7 (0.7%)
5	3 (0.3%)
6	1 (0.1%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func__ABS	
0	918 (92%)
1	67 (6.7%)
2	10 (1.0%)
3	2 (0.2%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__FLOOR	
0	895 (90%)
1	86 (8.6%)
2	16 (1.6%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func__TRUNC	
0	924 (92%)
1	63 (6.3%)
2	11 (1.1%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func__DAY	
0	875 (88%)
1	100 (10%)
2	16 (1.6%)
3	4 (0.4%)
4	4 (0.4%)
5	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0, 3, 68] 5 / 8
SELECT_n_of_aggr_func__AVG	
0	928 (93%)
1	47 (4.7%)
2	18 (1.8%)
3	6 (0.6%)
4	1 (0.1%)
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 14.00] 0.81 / 1.87
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 12.00] 0.74 / 1.77
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 13.00] 0.95 / 2.07
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 11.00] 0.86 / 1.92

Query parameter or value	Descriptive statistics
SELECT_n_of_aggr_func__SUM	
0	931 (93%)
1	46 (4.6%)
2	20 (2.0%)
3	2 (0.2%)
4	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.00, 3.00, 17.00] 3.56 / 3.08
FROM_n_of_join_paths	
1	163 (16%)
2	166 (17%)
3	176 (18%)
4	178 (18%)
5	165 (17%)
6	152 (15%)
FROM_n_of_super_joins__RIGHT	
0	553 (55%)
1	306 (31%)
2	119 (12%)
3	20 (2.0%)
4	2 (0.2%)
FROM_n_of_super_joins__FULL	
0	559 (56%)
1	302 (30%)
2	101 (10%)
3	32 (3.2%)
4	5 (0.5%)
5	1 (0.1%)
FROM_n_of_super_joins__LEFT	
0	540 (54%)
1	325 (33%)
2	102 (10%)
3	29 (2.9%)
4	4 (0.4%)
FROM_n_of_joins__INNER	[0.00, 3.00, 13.00] 3.22 / 2.48
FROM_n_of_joins__RIGHT	[0.00, 3.00, 13.00] 3.32 / 2.49
FROM_n_of_processed_rows	[5, 18,900,027,825, 88,440,166,902] 23,909,561,524 / 19,654,694,994
WHERE_n_of_predicates	[0.00, 3.00, 12.00] 3.70 / 2.85
WHERE_n_of_attribs_of_type__character	
0	666 (67%)
1	231 (23%)
2	79 (7.9%)
3	21 (2.1%)
4	3 (0.3%)
WHERE_n_of_attribs_of_type__character_varying	
0	377 (38%)

Query parameter or value	Descriptive statistics
1	306 (31%)
2	170 (17%)
3	96 (9.6%)
4	38 (3.8%)
5	9 (0.9%)
6	2 (0.2%)
7	2 (0.2%)
WHERE_n_of_attribs_of_type___integer	
0	365 (37%)
1	296 (30%)
2	193 (19%)
3	99 (9.9%)
4	33 (3.3%)
5	11 (1.1%)
6	2 (0.2%)
7	1 (0.1%)
WHERE_n_of_attribs_of_type___numeric	
0	565 (57%)
1	289 (29%)
2	117 (12%)
3	22 (2.2%)
4	7 (0.7%)
WHERE_n_of_attribs_of_type___date	
0	774 (77%)
1	180 (18%)
2	41 (4.1%)
3	3 (0.3%)
4	2 (0.2%)
WHERE_n_of_pkey_attribs	
0	542 (54%)
1	305 (31%)
2	110 (11%)
3	37 (3.7%)
4	5 (0.5%)
5	1 (0.1%)
WHERE_n_of_connect_OR	[0.00, 1.00, 11.00] 2.10 / 2.38
WHERE_n_of_operators___between	
0	548 (55%)
1	295 (30%)
2	110 (11%)
3	35 (3.5%)
4	8 (0.8%)
5	4 (0.4%)
WHERE_n_of_operators___greater_or_less	
0	329 (33%)
1	276 (28%)

Query parameter or value	Descriptive statistics
2	188 (19%)
3	114 (11%)
4	61 (6.1%)
5	22 (2.2%)
6	8 (0.8%)
7	2 (0.2%)
WHERE_n_of_operators_like	
0	787 (79%)
1	180 (18%)
2	31 (3.1%)
3	1 (0.1%)
4	1 (0.1%)
WHERE_n_of_operators_in	
0	567 (57%)
1	255 (26%)
2	117 (12%)
3	51 (5.1%)
4	10 (1.0%)
WHERE_n_of_non_aggr_func_TRUNC	
0	957 (96%)
1	43 (4.3%)
WHERE_n_of_non_aggr_func_SQRT	
0	874 (87%)
1	119 (12%)
2	6 (0.6%)
4	1 (0.1%)
WHERE_n_of_non_aggr_func_ABS	
0	956 (96%)
1	43 (4.3%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_DAY	
0	956 (96%)
1	43 (4.3%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_ROUND	
0	955 (96%)
1	44 (4.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_DOW	
0	959 (96%)
1	39 (3.9%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func_FLOOR	
0	962 (96%)
1	37 (3.7%)
2	1 (0.1%)

Query parameter or value	Descriptive statistics
WHERE_n_of_non_aggr_func__YEAR	
0	965 (97%)
1	34 (3.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__MONTH	
0	961 (96%)
1	39 (3.9%)
WHERE_n_of_non_aggr_func__LOG	
0	937 (94%)
1	61 (6.1%)
2	2 (0.2%)
WHERE_n_of_all_non_aggr_func	
0	620 (62%)
1	263 (26%)
2	89 (8.9%)
3	21 (2.1%)
4	3 (0.3%)
5	4 (0.4%)
GROUP_BY_n_of_columns	[0.00, 2.00, 12.00] 3.03 / 2.87
HAVING_n_of_main_predicates	
0	311 (31%)
1	84 (8.4%)
2	143 (14%)
3	186 (19%)
4	131 (13%)
5	88 (8.8%)
6	57 (5.7%)
HAVING_n_of_main_predicates__scalar_subquery	
0	464 (46%)
1	174 (17%)
2	174 (17%)
3	188 (19%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 25.0] 2.2 / 3.4
HAVING_n_of_processed_rows_by_subquery	[0, 300,000,018, 1,230,002,087,375] 51,915,009,209 / 107,768,543,734
ORDER_BY_n_of_columns	
0	299 (30%)
1	324 (32%)
2	239 (24%)
3	96 (9.6%)
4	27 (2.7%)
5	9 (0.9%)
6	3 (0.3%)
7	2 (0.2%)
9	1 (0.1%)
limit	[1, 488, 998] 489 / 285

Query parameter or value	Descriptive statistics
offset	[0, 0, 1,000] 238 / 319