

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

```
## Scenario: 1-3jp_ujpp__g__h0-3ss_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, 6, 99] 10 / 12
SELECT_n_of_non_aggr_func__DAY	
0	921 (92%)
1	66 (6.6%)
2	10 (1.0%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func__LOWER	
0	698 (70%)
1	235 (24%)
2	52 (5.2%)
3	11 (1.1%)
4	2 (0.2%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func__MONTH	
0	851 (85%)

Query parameter or value	Descriptive statistics
1	111 (11%)
2	19 (1.9%)
3	12 (1.2%)
4	5 (0.5%)
6	2 (0.2%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	677 (68%)
1	264 (26%)
2	45 (4.5%)
3	9 (0.9%)
4	3 (0.3%)
5	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___UPPER	
0	687 (69%)
1	238 (24%)
2	54 (5.4%)
3	16 (1.6%)
4	3 (0.3%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___LTRIM	
0	708 (71%)
1	225 (23%)
2	57 (5.7%)
3	6 (0.6%)
4	2 (0.2%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___RTRIM	
0	702 (70%)
1	224 (22%)
2	57 (5.7%)
3	14 (1.4%)
4	1 (0.1%)
6	2 (0.2%)
SELECT_n_of_non_aggr_func___SQRT	
0	835 (84%)
1	135 (14%)
2	20 (2.0%)
3	6 (0.6%)
4	3 (0.3%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___YEAR	
0	868 (87%)
1	102 (10%)
2	21 (2.1%)

Query parameter or value	Descriptive statistics
3	6 (0.6%)
4	1 (0.1%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___DOW	
0	887 (89%)
1	88 (8.8%)
2	19 (1.9%)
3	5 (0.5%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___FLOOR	
0	950 (95%)
1	43 (4.3%)
2	4 (0.4%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func___LOG	
0	910 (91%)
1	73 (7.3%)
2	11 (1.1%)
3	4 (0.4%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func___TRUNC	
0	948 (95%)
1	46 (4.6%)
2	6 (0.6%)
SELECT_n_of_non_aggr_func___ABS	
0	950 (95%)
1	44 (4.4%)
2	5 (0.5%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___ROUND	
0	959 (96%)
1	39 (3.9%)
2	2 (0.2%)
SELECT_n_of_all_non_aggr_func	[0.00, 2.00, 45.00] 3.16 / 4.17
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 10.00] 0.63 / 1.44
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 9.00] 0.66 / 1.48
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 11.00] 0.57 / 1.36
SELECT_n_of_aggr_func___MIN	
0	767 (77%)
1	89 (8.9%)
2	53 (5.3%)
3	31 (3.1%)
4	24 (2.4%)
5	21 (2.1%)
6	8 (0.8%)

Query parameter or value	Descriptive statistics
7	4 (0.4%)
8	3 (0.3%)
SELECT__n_of__aggr__func__AVG	
0	947 (95%)
1	37 (3.7%)
2	13 (1.3%)
3	2 (0.2%)
4	1 (0.1%)
SELECT__n_of__aggr__func__SUM	
0	929 (93%)
1	53 (5.3%)
2	13 (1.3%)
3	4 (0.4%)
4	1 (0.1%)
SELECT__n_of__all__aggr__func	[0.00, 2.00, 12.00] 2.62 / 2.25
FROM__n_of__join__paths	
1	354 (35%)
2	325 (33%)
3	321 (32%)
FROM__n_of__super__joins__FULL	
0	782 (78%)
1	192 (19%)
2	26 (2.6%)
FROM__n_of__super__joins__RIGHT	
0	779 (78%)
1	209 (21%)
2	12 (1.2%)
FROM__n_of__super__joins__LEFT	
0	780 (78%)
1	198 (20%)
2	22 (2.2%)
FROM__n_of__joins__INNER	
0	245 (25%)
1	267 (27%)
2	209 (21%)
3	147 (15%)
4	77 (7.7%)
5	30 (3.0%)
6	15 (1.5%)
7	8 (0.8%)
8	2 (0.2%)
FROM__n_of__joins__RIGHT	
0	275 (28%)
1	232 (23%)
2	202 (20%)
3	157 (16%)

Query parameter or value	Descriptive statistics
4	86 (8.6%)
5	27 (2.7%)
6	16 (1.6%)
7	4 (0.4%)
9	1 (0.1%)
FROM_n_of_processed_rows	[5, 6,809,989,734, 22,949,969,212] 6,505,781,715 / 6,070,782,126
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.79 / 2.21
WHERE_n_of_attribs_of_type__character_varying	
0	419 (42%)
1	339 (34%)
2	153 (15%)
3	73 (7.3%)
4	15 (1.5%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type__integer	
0	474 (47%)
1	311 (31%)
2	136 (14%)
3	58 (5.8%)
4	18 (1.8%)
5	2 (0.2%)
6	1 (0.1%)
WHERE_n_of_attribs_of_type__numeric	
0	626 (63%)
1	289 (29%)
2	66 (6.6%)
3	16 (1.6%)
4	2 (0.2%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	828 (83%)
1	150 (15%)
2	19 (1.9%)
3	3 (0.3%)
WHERE_n_of_attribs_of_type__character	
0	726 (73%)
1	219 (22%)
2	45 (4.5%)
3	9 (0.9%)
4	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	615 (62%)
1	286 (29%)
2	76 (7.6%)
3	21 (2.1%)

Query parameter or value	Descriptive statistics
4	2 (0.2%)
WHERE_n_of_connect_OR	[0.00, 1.00, 9.00] 1.38 / 1.67
WHERE_n_of_operators___greater_or_less	
0	401 (40%)
1	310 (31%)
2	177 (18%)
3	67 (6.7%)
4	34 (3.4%)
5	11 (1.1%)
WHERE_n_of_operators___like	
0	832 (83%)
1	145 (15%)
2	21 (2.1%)
3	2 (0.2%)
WHERE_n_of_operators___in	
0	629 (63%)
1	259 (26%)
2	94 (9.4%)
3	16 (1.6%)
5	2 (0.2%)
WHERE_n_of_operators___between	
0	609 (61%)
1	271 (27%)
2	93 (9.3%)
3	24 (2.4%)
4	3 (0.3%)
WHERE_n_of_non_aggr_func___DAY	
0	970 (97%)
1	30 (3.0%)
WHERE_n_of_non_aggr_func___MONTH	
0	968 (97%)
1	32 (3.2%)
WHERE_n_of_non_aggr_func___FLOOR	
0	973 (97%)
1	26 (2.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___SQRT	
0	895 (90%)
1	102 (10%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func___ROUND	
0	970 (97%)
1	30 (3.0%)
WHERE_n_of_non_aggr_func___ABS	
0	966 (97%)
1	34 (3.4%)

Query parameter or value	Descriptive statistics
WHERE_n_of_non_aggr_func__LOG	
0	944 (94%)
1	55 (5.5%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__TRUNC	
0	971 (97%)
1	29 (2.9%)
WHERE_n_of_non_aggr_func__DOW	
0	972 (97%)
1	28 (2.8%)
WHERE_n_of_non_aggr_func__YEAR	
0	975 (98%)
1	25 (2.5%)
WHERE_n_of_all_non_aggr_func	
0	675 (68%)
1	258 (26%)
2	59 (5.9%)
3	7 (0.7%)
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	297 (30%)
1	79 (7.9%)
2	135 (14%)
3	179 (18%)
4	153 (15%)
5	106 (11%)
6	51 (5.1%)
HAVING_n_of_main_predicates__scalar_subquery	
0	422 (42%)
1	187 (19%)
2	196 (20%)
3	195 (20%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 23.0] 2.5 / 3.6
HAVING_n_of_processed_rows_by_subqueries	[0, 540,000,090, 298,349,599,756] 16,268,589,484 / 34,982,718,592
ORDER_BY_n_of_columns	[0.00, 1.00, 9.00] 1.15 / 1.09
limit	[3, 495, 1,000] 501 / 293
offset	[0, 81, 997] 267 / 325