

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

## Scenario: 1-3jp\_ujpp\_\_g\_\_h0-3ss\_h0-3nss

## Scale factor: 50 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, 94] 9 / 11
SELECT_n_of_non_aggr_func__ABS	
0	952 (95%)
1	44 (4.4%)
2	2 (0.2%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func__DOW	
0	893 (89%)
1	92 (9.2%)
2	11 (1.1%)
3	2 (0.2%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__LTRIM	
0	694 (69%)

Query parameter or value	Descriptive statistics
1	238 (24%)
2	50 (5.0%)
3	15 (1.5%)
4	1 (0.1%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__ROUND	
0	946 (95%)
1	48 (4.8%)
2	5 (0.5%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	688 (69%)
1	230 (23%)
2	60 (6.0%)
3	15 (1.5%)
4	4 (0.4%)
5	3 (0.3%)
SELECT_n_of_non_aggr_func__UPPER	
0	706 (71%)
1	222 (22%)
2	54 (5.4%)
3	12 (1.2%)
4	3 (0.3%)
5	3 (0.3%)
SELECT_n_of_non_aggr_func__YEAR	
0	865 (87%)
1	116 (12%)
2	10 (1.0%)
3	6 (0.6%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__LOWER	
0	695 (70%)
1	228 (23%)
2	61 (6.1%)
3	10 (1.0%)
4	3 (0.3%)
5	2 (0.2%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__RTRIM	
0	683 (68%)
1	242 (24%)
2	61 (6.1%)
3	11 (1.1%)
4	2 (0.2%)

Query parameter or value	Descriptive statistics
5	1 (0.1%)
SELECT_n_of_non_aggr_func__MONTH	
0	837 (84%)
1	128 (13%)
2	26 (2.6%)
3	7 (0.7%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__SQRT	
0	865 (87%)
1	107 (11%)
2	18 (1.8%)
3	6 (0.6%)
4	4 (0.4%)
SELECT_n_of_non_aggr_func__DAY	
0	903 (90%)
1	90 (9.0%)
2	6 (0.6%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__TRUNC	
0	946 (95%)
1	50 (5.0%)
2	4 (0.4%)
SELECT_n_of_non_aggr_func__LOG	
0	901 (90%)
1	83 (8.3%)
2	14 (1.4%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func__FLOOR	
0	946 (95%)
1	50 (5.0%)
2	3 (0.3%)
3	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.0, 2.0, 32.0] 3.2 / 3.8
SELECT_n_of_aggr_func__AVG	
0	937 (94%)
1	52 (5.2%)
2	8 (0.8%)
3	2 (0.2%)
4	1 (0.1%)
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 9.00] 0.62 / 1.43
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 10.00] 0.62 / 1.46
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 10.00] 0.71 / 1.54
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 9.00] 0.63 / 1.48
SELECT_n_of_aggr_func__SUM	
0	946 (95%)
1	37 (3.7%)

Query parameter or value	Descriptive statistics
2	13 (1.3%)
3	4 (0.4%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 10.00] 2.73 / 2.29
FROM_n_of_join_paths	
1	326 (33%)
2	315 (32%)
3	359 (36%)
FROM_n_of_super_joins__FULL	
0	772 (77%)
1	203 (20%)
2	25 (2.5%)
FROM_n_of_super_joins__LEFT	
0	759 (76%)
1	221 (22%)
2	20 (2.0%)
FROM_n_of_super_joins__RIGHT	
0	771 (77%)
1	207 (21%)
2	22 (2.2%)
FROM_n_of_joins__INNER	
0	221 (22%)
1	288 (29%)
2	213 (21%)
3	149 (15%)
4	78 (7.8%)
5	32 (3.2%)
6	15 (1.5%)
7	3 (0.3%)
9	1 (0.1%)
FROM_n_of_joins__RIGHT	[0.00, 2.00, 11.00] 1.86 / 1.53
FROM_n_of_processed_rows	[5, 341,492,010, 1,140,049,222] 340,028,393 / 305,698,321
WHERE_n_of_predicates	[0.00, 3.00, 10.00] 2.86 / 2.21
WHERE_n_of_attribs_of_type__character	
0	715 (72%)
1	217 (22%)
2	56 (5.6%)
3	10 (1.0%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type__character_varying	
0	428 (43%)
1	322 (32%)
2	165 (17%)
3	66 (6.6%)
4	15 (1.5%)
5	4 (0.4%)

Query parameter or value	Descriptive statistics
WHERE_n_of_attribs_of_type__integer	
0	449 (45%)
1	309 (31%)
2	169 (17%)
3	58 (5.8%)
4	13 (1.3%)
5	2 (0.2%)
WHERE_n_of_attribs_of_type__numeric	
0	638 (64%)
1	262 (26%)
2	86 (8.6%)
3	12 (1.2%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type__date	
0	833 (83%)
1	138 (14%)
2	25 (2.5%)
3	3 (0.3%)
4	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	632 (63%)
1	264 (26%)
2	85 (8.5%)
3	17 (1.7%)
4	2 (0.2%)
WHERE_n_of_connect_OR	[0.00, 1.00, 9.00] 1.41 / 1.71
WHERE_n_of_operators__between	
0	610 (61%)
1	285 (29%)
2	81 (8.1%)
3	20 (2.0%)
4	2 (0.2%)
5	2 (0.2%)
WHERE_n_of_operators__greater_or_less	
0	402 (40%)
1	291 (29%)
2	166 (17%)
3	85 (8.5%)
4	43 (4.3%)
5	11 (1.1%)
6	2 (0.2%)
WHERE_n_of_operators__in	
0	612 (61%)
1	280 (28%)
2	93 (9.3%)
3	15 (1.5%)

Query parameter or value	Descriptive statistics
WHERE_n_of_operators_like	
0	840 (84%)
1	139 (14%)
2	21 (2.1%)
WHERE_n_of_non_aggr_func_sqrt	
0	896 (90%)
1	97 (9.7%)
2	7 (0.7%)
WHERE_n_of_non_aggr_func_month	
0	968 (97%)
1	32 (3.2%)
WHERE_n_of_non_aggr_func_dow	
0	966 (97%)
1	31 (3.1%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func_abs	
0	979 (98%)
1	20 (2.0%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_round	
0	975 (98%)
1	24 (2.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_floor	
0	975 (98%)
1	24 (2.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_day	
0	966 (97%)
1	32 (3.2%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func_log	
0	948 (95%)
1	51 (5.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_year	
0	974 (97%)
1	26 (2.6%)
WHERE_n_of_non_aggr_func_trunc	
0	967 (97%)
1	32 (3.2%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	691 (69%)
1	225 (23%)
2	75 (7.5%)

Query parameter or value	Descriptive statistics
3	8 (0.8%)
4	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.45 / 2.10
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.53 / 2.51
HAVING_n_of_main_predicates__non_scalar_subquery	
0	406 (41%)
1	209 (21%)
2	182 (18%)
3	203 (20%)
HAVING_n_of_main_predicates__scalar_subquery	
0	413 (41%)
1	189 (19%)
2	187 (19%)
3	211 (21%)
HAVING_n_of_subqueries__non_scalar_subquery	
0	406 (41%)
1	209 (21%)
2	182 (18%)
3	203 (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, 359,940,864, 14,820,639,821] 1,297,693,643 / 2,158,605,956
ORDER_BY_n_of_columns	
0	306 (31%)
1	406 (41%)
2	222 (22%)
3	50 (5.0%)
4	10 (1.0%)
5	4 (0.4%)
6	1 (0.1%)
7	1 (0.1%)
limit	[2, 526, 1,000] 515 / 289
offset	[0, 0, 999] 242 / 317