

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

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

## Scale factor: 100 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, 92] 10 / 12
SELECT_n_of_non_aggr_func__FLOOR	
0	938 (94%)
1	55 (5.5%)
2	7 (0.7%)
SELECT_n_of_non_aggr_func__LOG	
0	909 (91%)
1	79 (7.9%)
2	10 (1.0%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func__RTRIM	
0	680 (68%)
1	245 (25%)
2	57 (5.7%)
3	12 (1.2%)

Query parameter or value	Descriptive statistics
4	5 (0.5%)
5	1 (0.1%)
SELECT__n_of__non__aggr__func__UPPER	
0	680 (68%)
1	259 (26%)
2	48 (4.8%)
3	11 (1.1%)
4	2 (0.2%)
SELECT__n_of__non__aggr__func__SQRT	
0	865 (87%)
1	103 (10%)
2	22 (2.2%)
3	9 (0.9%)
5	1 (0.1%)
SELECT__n_of__non__aggr__func__DAY	
0	923 (92%)
1	69 (6.9%)
2	5 (0.5%)
3	3 (0.3%)
SELECT__n_of__non__aggr__func__LTRIM	
0	675 (68%)
1	252 (25%)
2	59 (5.9%)
3	8 (0.8%)
4	2 (0.2%)
5	3 (0.3%)
6	1 (0.1%)
SELECT__n_of__non__aggr__func__MONTH	
0	868 (87%)
1	100 (10%)
2	22 (2.2%)
3	6 (0.6%)
4	2 (0.2%)
6	1 (0.1%)
7	1 (0.1%)
SELECT__n_of__non__aggr__func__SUBSTR	
0	664 (66%)
1	279 (28%)
2	44 (4.4%)
3	9 (0.9%)
4	2 (0.2%)
5	2 (0.2%)
SELECT__n_of__non__aggr__func__ABS	
0	944 (94%)
1	53 (5.3%)
2	3 (0.3%)

Query parameter or value	Descriptive statistics
SELECT_n_of_non_aggr_func__LOWER	
0	711 (71%)
1	225 (23%)
2	48 (4.8%)
3	10 (1.0%)
4	6 (0.6%)
SELECT_n_of_non_aggr_func__TRUNC	
0	943 (94%)
1	53 (5.3%)
2	4 (0.4%)
SELECT_n_of_non_aggr_func__YEAR	
0	889 (89%)
1	93 (9.3%)
2	13 (1.3%)
3	3 (0.3%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__DOW	
0	901 (90%)
1	83 (8.3%)
2	12 (1.2%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__ROUND	
0	954 (95%)
1	39 (3.9%)
2	5 (0.5%)
3	2 (0.2%)
SELECT_n_of_all_non_aggr_func	[0.00, 2.00, 32.00] 3.08 / 3.74
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 9.00] 0.69 / 1.54
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 10.00] 0.72 / 1.54
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 9.00] 0.62 / 1.44
SELECT_n_of_aggr_func__COUNT_DISTINCT	
0	758 (76%)
1	93 (9.3%)
2	50 (5.0%)
3	37 (3.7%)
4	28 (2.8%)
5	17 (1.7%)
6	12 (1.2%)
7	2 (0.2%)
8	3 (0.3%)
SELECT_n_of_aggr_func__SUM	
0	939 (94%)
1	45 (4.5%)
2	15 (1.5%)

Query parameter or value	Descriptive statistics
3	1 (0.1%)
SELECT_n_of_aggr_func__AVG	
0	945 (95%)
1	41 (4.1%)
2	14 (1.4%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 10.00] 2.78 / 2.27
FROM_n_of_join_paths	
1	335 (34%)
2	338 (34%)
3	327 (33%)
FROM_n_of_super_joins__LEFT	
0	764 (76%)
1	214 (21%)
2	22 (2.2%)
FROM_n_of_super_joins__RIGHT	
0	792 (79%)
1	191 (19%)
2	17 (1.7%)
FROM_n_of_super_joins__FULL	
0	779 (78%)
1	197 (20%)
2	24 (2.4%)
FROM_n_of_joins__INNER	[0.00, 2.00, 9.00] 1.80 / 1.55
FROM_n_of_joins__RIGHT	[0.00, 1.00, 10.00] 1.73 / 1.53
FROM_n_of_processed_rows	[5, 681,037,927, 2,295,113,791] 635,746,785 / 593,032,345
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.75 / 2.17
WHERE_n_of_attribs_of_type__character_varying	
0	434 (43%)
1	319 (32%)
2	167 (17%)
3	53 (5.3%)
4	24 (2.4%)
5	2 (0.2%)
6	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	831 (83%)
1	147 (15%)
2	20 (2.0%)
3	2 (0.2%)
WHERE_n_of_attribs_of_type__integer	
0	442 (44%)
1	343 (34%)
2	140 (14%)
3	59 (5.9%)
4	14 (1.4%)

Query parameter or value	Descriptive statistics
5	2 (0.2%)
WHERE_n_of_attribs_of_type__numeric	
0	667 (67%)
1	246 (25%)
2	72 (7.2%)
3	15 (1.5%)
WHERE_n_of_attribs_of_type__character	
0	729 (73%)
1	219 (22%)
2	45 (4.5%)
3	7 (0.7%)
WHERE_n_of_pkey_attribs	
0	612 (61%)
1	293 (29%)
2	73 (7.3%)
3	20 (2.0%)
4	2 (0.2%)
WHERE_n_of_connect_OR	
0	453 (45%)
1	193 (19%)
2	132 (13%)
3	91 (9.1%)
4	75 (7.5%)
5	29 (2.9%)
6	19 (1.9%)
7	7 (0.7%)
8	1 (0.1%)
WHERE_n_of_operators__between	
0	633 (63%)
1	250 (25%)
2	94 (9.4%)
3	17 (1.7%)
4	6 (0.6%)
WHERE_n_of_operators__greater_or_less	
0	426 (43%)
1	319 (32%)
2	156 (16%)
3	61 (6.1%)
4	29 (2.9%)
5	7 (0.7%)
6	1 (0.1%)
7	1 (0.1%)
WHERE_n_of_operators__like	
0	822 (82%)
1	157 (16%)
2	18 (1.8%)

Query parameter or value	Descriptive statistics
3	3 (0.3%)
WHERE_n_of_operators_in	
0	591 (59%)
1	302 (30%)
2	78 (7.8%)
3	24 (2.4%)
4	5 (0.5%)
WHERE_n_of_non_aggr_func_DOW	
0	985 (99%)
1	14 (1.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_SQRT	
0	923 (92%)
1	72 (7.2%)
2	4 (0.4%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func_TRUNC	
0	970 (97%)
1	30 (3.0%)
WHERE_n_of_non_aggr_func_ABS	
0	967 (97%)
1	32 (3.2%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_MONTH	
0	968 (97%)
1	32 (3.2%)
WHERE_n_of_non_aggr_func_YEAR	
0	970 (97%)
1	30 (3.0%)
WHERE_n_of_non_aggr_func_LOG	
0	936 (94%)
1	60 (6.0%)
2	4 (0.4%)
WHERE_n_of_non_aggr_func_DAY	
0	980 (98%)
1	20 (2.0%)
WHERE_n_of_non_aggr_func_FLOOR	
0	977 (98%)
1	22 (2.2%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_ROUND	
0	973 (97%)
1	27 (2.7%)
WHERE_n_of_all_non_aggr_func	
0	713 (71%)
1	222 (22%)

Query parameter or value	Descriptive statistics
2	53 (5.3%)
3	12 (1.2%)
GROUP_BY_n_of_columns	[0.00, 2.00, 11.00] 2.32 / 2.11
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.46 / 2.57
HAVING_n_of_main_predicates__non_scalar_subquery	
0	437 (44%)
1	183 (18%)
2	166 (17%)
3	214 (21%)
HAVING_n_of_main_predicates__scalar_subquery	
0	416 (42%)
1	206 (21%)
2	192 (19%)
3	186 (19%)
HAVING_n_of_subqueries__non_scalar_subquery	
0	437 (44%)
1	183 (18%)
2	166 (17%)
3	214 (21%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 24.0] 2.4 / 3.6
HAVING_n_of_processed_rows_by_subqueries	[0, 412,500,313, 34,051,895,100] 2,149,160,087 / 3,728,783,720
ORDER_BY_n_of_columns	
0	309 (31%)
1	388 (39%)
2	236 (24%)
3	47 (4.7%)
4	12 (1.2%)
5	6 (0.6%)
6	2 (0.2%)
limit	[1, 491, 1,000] 500 / 290
offset	[0, 0, 998] 255 / 323