

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

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

## Scale factor: 100 GB

## The 1000-query pack was generated on: 2023-11-06

## 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, 102] 10 / 14
SELECT_n_of_non_aggr_func__RTRIM	
0	721 (72%)
1	206 (21%)
2	51 (5.1%)
3	14 (1.4%)
4	6 (0.6%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	690 (69%)
1	240 (24%)
2	54 (5.4%)
3	8 (0.8%)
4	3 (0.3%)

Query parameter or value	Descriptive statistics
5	3 (0.3%)
6	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__LOG	
0	900 (90%)
1	87 (8.7%)
2	9 (0.9%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__LOWER	
0	699 (70%)
1	227 (23%)
2	55 (5.5%)
3	16 (1.6%)
4	3 (0.3%)
SELECT_n_of_non_aggr_func__LTRIM	
0	718 (72%)
1	210 (21%)
2	52 (5.2%)
3	14 (1.4%)
4	3 (0.3%)
5	3 (0.3%)
SELECT_n_of_non_aggr_func__MONTH	
0	848 (85%)
1	117 (12%)
2	21 (2.1%)
3	7 (0.7%)
4	3 (0.3%)
5	3 (0.3%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__SQRT	
0	850 (85%)
1	120 (12%)
2	20 (2.0%)
3	4 (0.4%)
4	5 (0.5%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__UPPER	
0	694 (69%)
1	223 (22%)
2	61 (6.1%)
3	18 (1.8%)
4	2 (0.2%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__FLOOR	

Query parameter or value	Descriptive statistics
0	945 (95%)
1	52 (5.2%)
2	1 (0.1%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___ABS	
0	944 (94%)
1	50 (5.0%)
2	5 (0.5%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___DOW	
0	880 (88%)
1	96 (9.6%)
2	10 (1.0%)
3	6 (0.6%)
4	6 (0.6%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___ROUND	
0	944 (94%)
1	54 (5.4%)
2	2 (0.2%)
SELECT_n_of_non_aggr_func___YEAR	
0	865 (87%)
1	106 (11%)
2	19 (1.9%)
3	4 (0.4%)
4	5 (0.5%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___TRUNC	
0	945 (95%)
1	52 (5.2%)
2	3 (0.3%)
SELECT_n_of_non_aggr_func___DAY	
0	921 (92%)
1	67 (6.7%)
2	10 (1.0%)
4	2 (0.2%)
SELECT_n_of_all_non_aggr_func	[0.0, 2.0, 38.0] 3.2 / 4.5
SELECT_n_of_aggr_func___AVG	
0	926 (93%)
1	52 (5.2%)
2	18 (1.8%)
3	4 (0.4%)
SELECT_n_of_aggr_func___MAX	
0	740 (74%)

Query parameter or value	Descriptive statistics
1	81 (8.1%)
2	71 (7.1%)
3	34 (3.4%)
4	31 (3.1%)
5	22 (2.2%)
6	14 (1.4%)
7	6 (0.6%)
8	1 (0.1%)
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 11.00] 0.61 / 1.42
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 10.00] 0.59 / 1.41
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 11.00] 0.65 / 1.44
SELECT_n_of_aggr_func__SUM	
0	940 (94%)
1	47 (4.7%)
2	13 (1.3%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 11.00] 2.71 / 2.27
FROM_n_of_join_paths	
1	345 (35%)
2	327 (33%)
3	328 (33%)
FROM_n_of_super_joins__RIGHT	
0	770 (77%)
1	212 (21%)
2	18 (1.8%)
FROM_n_of_super_joins__LEFT	
0	803 (80%)
1	178 (18%)
2	19 (1.9%)
FROM_n_of_super_joins__FULL	
0	781 (78%)
1	195 (20%)
2	24 (2.4%)
FROM_n_of_joins__INNER	
0	272 (27%)
1	250 (25%)
2	222 (22%)
3	121 (12%)
4	73 (7.3%)
5	42 (4.2%)
6	10 (1.0%)
7	6 (0.6%)
8	4 (0.4%)
FROM_n_of_joins__RIGHT	
0	243 (24%)
1	269 (27%)
2	224 (22%)

Query parameter or value	Descriptive statistics
3	138 (14%)
4	75 (7.5%)
5	25 (2.5%)
6	16 (1.6%)
7	8 (0.8%)
8	2 (0.2%)
FROM_n_of_processed_rows	[5, 681,037,927, 2,295,113,786] 647,545,503 / 600,842,516
WHERE_n_of_predicates	[0.00, 3.00, 10.00] 2.82 / 2.20
WHERE_n_of_attribs_of_type__character_varying	
0	430 (43%)
1	325 (33%)
2	154 (15%)
3	70 (7.0%)
4	18 (1.8%)
5	3 (0.3%)
WHERE_n_of_attribs_of_type__integer	
0	443 (44%)
1	304 (30%)
2	185 (19%)
3	48 (4.8%)
4	17 (1.7%)
5	3 (0.3%)
WHERE_n_of_attribs_of_type__numeric	
0	649 (65%)
1	257 (26%)
2	79 (7.9%)
3	14 (1.4%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	846 (85%)
1	126 (13%)
2	26 (2.6%)
3	2 (0.2%)
WHERE_n_of_attribs_of_type__character	
0	724 (72%)
1	222 (22%)
2	43 (4.3%)
3	10 (1.0%)
4	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	606 (61%)
1	282 (28%)
2	101 (10%)
3	9 (0.9%)
4	2 (0.2%)

Query parameter or value	Descriptive statistics
WHERE_n_of_connect_OR	
0	455 (46%)
1	167 (17%)
2	140 (14%)
3	114 (11%)
4	65 (6.5%)
5	30 (3.0%)
6	23 (2.3%)
7	5 (0.5%)
8	1 (0.1%)
WHERE_n_of_operators___greater_or_less	
0	418 (42%)
1	298 (30%)
2	162 (16%)
3	74 (7.4%)
4	34 (3.4%)
5	11 (1.1%)
6	2 (0.2%)
7	1 (0.1%)
WHERE_n_of_operators___between	
0	609 (61%)
1	278 (28%)
2	94 (9.4%)
3	17 (1.7%)
4	2 (0.2%)
WHERE_n_of_operators___in	
0	603 (60%)
1	294 (29%)
2	81 (8.1%)
3	19 (1.9%)
4	3 (0.3%)
WHERE_n_of_operators___like	
0	819 (82%)
1	163 (16%)
2	17 (1.7%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func___FLOOR	
0	963 (96%)
1	36 (3.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___LOG	
0	946 (95%)
1	54 (5.4%)
WHERE_n_of_non_aggr_func___SQRT	
0	913 (91%)
1	80 (8.0%)

Query parameter or value	Descriptive statistics
2	7 (0.7%)
WHERE_n_of_non_aggr_func__YEAR	
0	976 (98%)
1	24 (2.4%)
WHERE_n_of_non_aggr_func__DOW	
0	977 (98%)
1	22 (2.2%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__ABS	
0	974 (97%)
1	25 (2.5%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__ROUND	
0	974 (97%)
1	25 (2.5%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__MONTH	
0	970 (97%)
1	29 (2.9%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DAY	
0	980 (98%)
1	20 (2.0%)
WHERE_n_of_non_aggr_func__TRUNC	
0	977 (98%)
1	22 (2.2%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	715 (72%)
1	222 (22%)
2	51 (5.1%)
3	9 (0.9%)
4	3 (0.3%)
GROUP_BY_n_of_columns	
0	262 (26%)
1	186 (19%)
2	151 (15%)
3	135 (14%)
4	108 (11%)
5	79 (7.9%)
6	47 (4.7%)
7	22 (2.2%)
8	10 (1.0%)
HAVING_n_of_main_predicates	[0.00, 3.00, 9.00] 3.28 / 2.63
HAVING_n_of_main_predicates__scalar_subquery	
0	477 (48%)

Query parameter or value	Descriptive statistics
<i>1</i>	171 (17%)
<i>2</i>	175 (18%)
<i>3</i>	177 (18%)
HAVING_n_of_main_predicates__non_scalar_subquery	
<i>0</i>	469 (47%)
<i>1</i>	171 (17%)
<i>2</i>	162 (16%)
<i>3</i>	198 (20%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 24.0] 2.3 / 3.5
HAVING_n_of_subqueries__non_scalar_subquery	
<i>0</i>	469 (47%)
<i>1</i>	171 (17%)
<i>2</i>	162 (16%)
<i>3</i>	198 (20%)
HAVING_n_of_processed_rows_by_subqueries	[0, 361,000,150, 33,777,819,296] 2,145,162,434 / 3,979,348,136
ORDER_BY_n_of_columns	
<i>0</i>	306 (31%)
<i>1</i>	368 (37%)
<i>2</i>	247 (25%)
<i>3</i>	54 (5.4%)
<i>4</i>	12 (1.2%)
<i>5</i>	7 (0.7%)
<i>6</i>	4 (0.4%)
<i>7</i>	2 (0.2%)
limit	[2, 481, 1,000] 490 / 291
offset	[0, 14, 1,000] 254 / 328