

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

## Scenario: 1-1jp\_\_g\_\_h0ss\_h0nss

## Scale factor: 50 GB

## The 1000-query pack was generated on: 2023-11-04

## 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	[1.0, 4.0, 40.0] 5.6 / 5.0
SELECT_n_of_non_aggr_func__RTRIM	
0	771 (77%)
1	193 (19%)
2	32 (3.2%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	787 (79%)
1	190 (19%)
2	22 (2.2%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__LOWER	
0	809 (81%)
1	161 (16%)

Query parameter or value	Descriptive statistics
2	28 (2.8%)
3	2 (0.2%)
SELECT__n_of__non__aggr__func__LTRIM	
0	789 (79%)
1	184 (18%)
2	23 (2.3%)
3	4 (0.4%)
SELECT__n_of__non__aggr__func__UPPER	
0	792 (79%)
1	191 (19%)
2	14 (1.4%)
3	3 (0.3%)
SELECT__n_of__non__aggr__func__DOW	
0	943 (94%)
1	51 (5.1%)
2	5 (0.5%)
3	1 (0.1%)
SELECT__n_of__non__aggr__func__SQRT	
0	909 (91%)
1	83 (8.3%)
2	6 (0.6%)
3	1 (0.1%)
4	1 (0.1%)
SELECT__n_of__non__aggr__func__TRUNC	
0	966 (97%)
1	33 (3.3%)
2	1 (0.1%)
SELECT__n_of__non__aggr__func__YEAR	
0	921 (92%)
1	72 (7.2%)
2	7 (0.7%)
SELECT__n_of__non__aggr__func__LOG	
0	957 (96%)
1	40 (4.0%)
2	3 (0.3%)
SELECT__n_of__non__aggr__func__MONTH	
0	903 (90%)
1	84 (8.4%)
2	12 (1.2%)
3	1 (0.1%)
SELECT__n_of__non__aggr__func__FLOOR	
0	960 (96%)
1	40 (4.0%)
SELECT__n_of__non__aggr__func__ROUND	
0	953 (95%)
1	47 (4.7%)

Query parameter or value	Descriptive statistics
SELECT__n_of__non_aggr_func__DAY	
0	945 (95%)
1	48 (4.8%)
2	7 (0.7%)
SELECT__n_of__non_aggr_func__ABS	
0	979 (98%)
1	21 (2.1%)
SELECT__n_of__all__non_aggr_func	[0.00, 1.00, 16.00] 1.82 / 1.91
SELECT__n_of__aggr_func__COUNT_DISTINCT	
0	791 (79%)
1	112 (11%)
2	48 (4.8%)
3	30 (3.0%)
4	14 (1.4%)
5	3 (0.3%)
6	2 (0.2%)
SELECT__n_of__aggr_func__MIN	
0	761 (76%)
1	124 (12%)
2	65 (6.5%)
3	33 (3.3%)
4	10 (1.0%)
5	2 (0.2%)
6	5 (0.5%)
SELECT__n_of__aggr_func__COUNT	
0	767 (77%)
1	116 (12%)
2	60 (6.0%)
3	37 (3.7%)
4	12 (1.2%)
5	5 (0.5%)
6	2 (0.2%)
7	1 (0.1%)
SELECT__n_of__aggr_func__MAX	
0	792 (79%)
1	108 (11%)
2	60 (6.0%)
3	15 (1.5%)
4	15 (1.5%)
5	6 (0.6%)
6	1 (0.1%)
7	3 (0.3%)
SELECT__n_of__aggr_func__SUM	
0	956 (96%)
1	38 (3.8%)
2	6 (0.6%)

Query parameter or value	Descriptive statistics
SELECT__n_of__aggr_func__AVG	
0	947 (95%)
1	46 (4.6%)
2	5 (0.5%)
3	2 (0.2%)
SELECT__n_of__all_aggr_func	
0	213 (21%)
1	290 (29%)
2	228 (23%)
3	144 (14%)
4	74 (7.4%)
5	32 (3.2%)
6	12 (1.2%)
7	7 (0.7%)
FROM__n_of__join_paths	
1	1,000 (100%)
FROM__n_of__joins__INNER	
0	458 (46%)
1	334 (33%)
2	161 (16%)
3	41 (4.1%)
4	6 (0.6%)
FROM__n_of__joins__RIGHT	
0	527 (53%)
1	322 (32%)
2	122 (12%)
3	27 (2.7%)
4	2 (0.2%)
FROM__n_of__processed_rows	[5, 40,490,376, 382,516,427] 128,798,705 / 152,978,113
WHERE__n_of__predicates	
0	227 (23%)
1	254 (25%)
2	244 (24%)
3	143 (14%)
4	87 (8.7%)
5	36 (3.6%)
6	9 (0.9%)
WHERE__n_of__attribs_of_type__character_varying	
0	541 (54%)
1	332 (33%)
2	101 (10%)
3	21 (2.1%)
4	4 (0.4%)
5	1 (0.1%)
WHERE__n_of__attribs_of_type__integer	

Query parameter or value	Descriptive statistics
0	563 (56%)
1	319 (32%)
2	102 (10%)
3	16 (1.6%)
WHERE_n_of_attribs_of_type__numeric	
0	759 (76%)
1	196 (20%)
2	41 (4.1%)
3	4 (0.4%)
WHERE_n_of_attribs_of_type__character	
0	845 (85%)
1	143 (14%)
2	11 (1.1%)
3	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	901 (90%)
1	92 (9.2%)
2	7 (0.7%)
WHERE_n_of_pkey_attribs	
0	708 (71%)
1	255 (26%)
2	35 (3.5%)
3	2 (0.2%)
WHERE_n_of_connect_OR	
0	636 (64%)
1	211 (21%)
2	91 (9.1%)
3	51 (5.1%)
4	7 (0.7%)
5	4 (0.4%)
WHERE_n_of_operators__greater_or_less	
0	553 (55%)
1	286 (29%)
2	119 (12%)
3	34 (3.4%)
4	8 (0.8%)
WHERE_n_of_operators__in	
0	730 (73%)
1	229 (23%)
2	37 (3.7%)
3	4 (0.4%)
WHERE_n_of_operators__between	
0	728 (73%)
1	221 (22%)
2	44 (4.4%)
3	6 (0.6%)

Query parameter or value	Descriptive statistics
5	1 (0.1%)
WHERE_n_of_operators_like	
0	886 (89%)
1	111 (11%)
2	2 (0.2%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func__SQRT	
0	939 (94%)
1	61 (6.1%)
WHERE_n_of_non_aggr_func__LOG	
0	958 (96%)
1	41 (4.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__ROUND	
0	987 (99%)
1	12 (1.2%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DAY	
0	981 (98%)
1	19 (1.9%)
WHERE_n_of_non_aggr_func__TRUNC	
0	980 (98%)
1	20 (2.0%)
WHERE_n_of_non_aggr_func__YEAR	
0	993 (99%)
1	7 (0.7%)
WHERE_n_of_non_aggr_func__ABS	
0	986 (99%)
1	14 (1.4%)
WHERE_n_of_non_aggr_func__FLOOR	
0	982 (98%)
1	18 (1.8%)
WHERE_n_of_non_aggr_func__MONTH	
0	986 (99%)
1	14 (1.4%)
WHERE_n_of_non_aggr_func__DOW	
0	986 (99%)
1	14 (1.4%)
WHERE_n_of_all_non_aggr_func	
0	801 (80%)
1	177 (18%)
2	20 (2.0%)
3	1 (0.1%)
4	1 (0.1%)
GROUP_BY_n_of_columns	
0	268 (27%)

Query parameter or value	Descriptive statistics
<i>1</i>	258 (26%)
<i>2</i>	207 (21%)
<i>3</i>	164 (16%)
<i>4</i>	68 (6.8%)
<i>5</i>	28 (2.8%)
<i>6</i>	7 (0.7%)
HAVING_n_of_main_predicates	
<i>0</i>	450 (45%)
<i>1</i>	161 (16%)
<i>2</i>	193 (19%)
<i>3</i>	196 (20%)
ORDER_BY_n_of_columns	
<i>0</i>	284 (28%)
<i>1</i>	470 (47%)
<i>2</i>	212 (21%)
<i>3</i>	24 (2.4%)
<i>4</i>	8 (0.8%)
<i>5</i>	2 (0.2%)
limit	[1, 522, 1,000] 512 / 289
offset	[0, 0, 1,000] 242 / 323