

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

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

## Scale factor: 1 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] 5.9 / 5.7
SELECT_n_of_non_aggr_func__DAY	
0	934 (93%)
1	62 (6.2%)
2	4 (0.4%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	755 (76%)
1	214 (21%)
2	28 (2.8%)
3	2 (0.2%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__RTRIM	
0	782 (78%)
1	192 (19%)
2	22 (2.2%)

Query parameter or value	Descriptive statistics
3	4 (0.4%)
SELECT_n_of_non_aggr_func___LTRIM	
0	785 (79%)
1	186 (19%)
2	26 (2.6%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func___DOW	
0	920 (92%)
1	73 (7.3%)
2	6 (0.6%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___LOG	
0	932 (93%)
1	59 (5.9%)
2	8 (0.8%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___LOWER	
0	785 (79%)
1	183 (18%)
2	29 (2.9%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func___MONTH	
0	884 (88%)
1	98 (9.8%)
2	15 (1.5%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func___YEAR	
0	894 (89%)
1	93 (9.3%)
2	11 (1.1%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___UPPER	
0	791 (79%)
1	172 (17%)
2	32 (3.2%)
3	5 (0.5%)
SELECT_n_of_non_aggr_func___TRUNC	
0	968 (97%)
1	30 (3.0%)
2	2 (0.2%)
SELECT_n_of_non_aggr_func___ROUND	
0	960 (96%)
1	36 (3.6%)
2	4 (0.4%)
SELECT_n_of_non_aggr_func___SQRT	
0	894 (89%)

Query parameter or value	Descriptive statistics
1	96 (9.6%)
2	9 (0.9%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___ABS	
0	961 (96%)
1	39 (3.9%)
SELECT_n_of_non_aggr_func___FLOOR	
0	974 (97%)
1	24 (2.4%)
2	2 (0.2%)
SELECT_n_of_all_non_aggr_func	[0.00, 2.00, 17.00] 2.03 / 2.29
SELECT_n_of_aggr_func___COUNT	
0	793 (79%)
1	108 (11%)
2	53 (5.3%)
3	26 (2.6%)
4	8 (0.8%)
5	7 (0.7%)
6	4 (0.4%)
7	1 (0.1%)
SELECT_n_of_aggr_func___MIN	
0	740 (74%)
1	135 (14%)
2	68 (6.8%)
3	32 (3.2%)
4	17 (1.7%)
5	5 (0.5%)
6	2 (0.2%)
8	1 (0.1%)
SELECT_n_of_aggr_func___MAX	
0	778 (78%)
1	116 (12%)
2	52 (5.2%)
3	23 (2.3%)
4	20 (2.0%)
5	7 (0.7%)
6	4 (0.4%)
SELECT_n_of_aggr_func___COUNT_DISTINCT	
0	776 (78%)
1	110 (11%)
2	63 (6.3%)
3	28 (2.8%)
4	16 (1.6%)
5	3 (0.3%)
6	2 (0.2%)
7	1 (0.1%)

Query parameter or value	Descriptive statistics
8	1 (0.1%)
SELECT_n_of_aggr_func__SUM	
0	953 (95%)
1	40 (4.0%)
2	6 (0.6%)
3	1 (0.1%)
SELECT_n_of_aggr_func__AVG	
0	952 (95%)
1	43 (4.3%)
2	4 (0.4%)
3	1 (0.1%)
SELECT_n_of_all_aggr_func	
0	189 (19%)
1	306 (31%)
2	224 (22%)
3	144 (14%)
4	73 (7.3%)
5	39 (3.9%)
6	18 (1.8%)
7	4 (0.4%)
8	3 (0.3%)
FROM_n_of_join_paths	
1	1,000 (100%)
FROM_n_of_joins__INNER	
0	505 (51%)
1	320 (32%)
2	139 (14%)
3	34 (3.4%)
4	2 (0.2%)
FROM_n_of_joins__RIGHT	
0	484 (48%)
1	322 (32%)
2	151 (15%)
3	36 (3.6%)
4	7 (0.7%)
FROM_n_of_processed_rows	[5, 1,500,000, 7,650,621] 2,967,578 / 3,156,998
WHERE_n_of_predicates	
0	225 (23%)
1	227 (23%)
2	223 (22%)
3	179 (18%)
4	94 (9.4%)
5	40 (4.0%)
6	12 (1.2%)
WHERE_n_of_attribs_of_type__character	
0	806 (81%)

Query parameter or value	Descriptive statistics
1	168 (17%)
2	23 (2.3%)
3	3 (0.3%)
WHERE_n_of_attribs_of_type__date	
0	895 (90%)
1	92 (9.2%)
2	13 (1.3%)
WHERE_n_of_attribs_of_type__character_varying	
0	538 (54%)
1	323 (32%)
2	119 (12%)
3	18 (1.8%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type__integer	
0	542 (54%)
1	337 (34%)
2	103 (10%)
3	16 (1.6%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type__numeric	
0	745 (75%)
1	219 (22%)
2	32 (3.2%)
3	4 (0.4%)
WHERE_n_of_pkey_attribs	
0	700 (70%)
1	263 (26%)
2	34 (3.4%)
3	3 (0.3%)
WHERE_n_of_connect_OR	
0	591 (59%)
1	230 (23%)
2	104 (10%)
3	58 (5.8%)
4	14 (1.4%)
5	3 (0.3%)
WHERE_n_of_operators__greater_or_less	
0	519 (52%)
1	301 (30%)
2	139 (14%)
3	36 (3.6%)
4	3 (0.3%)
5	2 (0.2%)
WHERE_n_of_operators__in	
0	722 (72%)
1	229 (23%)

Query parameter or value	Descriptive statistics
2	40 (4.0%)
3	9 (0.9%)
WHERE_n_of_operators_between	
0	699 (70%)
1	239 (24%)
2	51 (5.1%)
3	10 (1.0%)
4	1 (0.1%)
WHERE_n_of_operators_like	
0	887 (89%)
1	109 (11%)
2	4 (0.4%)
WHERE_n_of_non_aggr_func_YEAR	
0	984 (98%)
1	16 (1.6%)
WHERE_n_of_non_aggr_func_LOG	
0	954 (95%)
1	44 (4.4%)
2	1 (0.1%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func_SQRT	
0	925 (93%)
1	71 (7.1%)
2	4 (0.4%)
WHERE_n_of_non_aggr_func_FLOOR	
0	988 (99%)
1	12 (1.2%)
WHERE_n_of_non_aggr_func_ABS	
0	976 (98%)
1	23 (2.3%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_ROUND	
0	982 (98%)
1	18 (1.8%)
WHERE_n_of_non_aggr_func_MONTH	
0	982 (98%)
1	18 (1.8%)
WHERE_n_of_non_aggr_func_DAY	
0	986 (99%)
1	14 (1.4%)
WHERE_n_of_non_aggr_func_DOW	
0	986 (99%)
1	13 (1.3%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_TRUNC	
0	986 (99%)

Query parameter or value	Descriptive statistics
<i>1</i>	14 (1.4%)
WHERE_n_of_all_non_aggr_func	
<i>0</i>	781 (78%)
<i>1</i>	184 (18%)
<i>2</i>	29 (2.9%)
<i>3</i>	6 (0.6%)
GROUP_BY_n_of_columns	
<i>0</i>	250 (25%)
<i>1</i>	264 (26%)
<i>2</i>	217 (22%)
<i>3</i>	136 (14%)
<i>4</i>	85 (8.5%)
<i>5</i>	42 (4.2%)
<i>6</i>	6 (0.6%)
HAVING_n_of_main_predicates	
<i>0</i>	447 (45%)
<i>1</i>	173 (17%)
<i>2</i>	182 (18%)
<i>3</i>	198 (20%)
ORDER_BY_n_of_columns	
<i>0</i>	311 (31%)
<i>1</i>	455 (46%)
<i>2</i>	195 (20%)
<i>3</i>	30 (3.0%)
<i>4</i>	7 (0.7%)
<i>5</i>	2 (0.2%)
limit	[1, 529, 1,000] 520 / 283
offset	[0, 0, 999] 249 / 325