

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

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

## Scale factor: 500 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, 41.0] 5.8 / 5.5
SELECT_n_of_non_aggr_func__LTRIM	
0	774 (77%)
1	196 (20%)
2	27 (2.7%)
3	2 (0.2%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__ABS	
0	963 (96%)
1	35 (3.5%)
2	2 (0.2%)
SELECT_n_of_non_aggr_func__SQRT	
0	913 (91%)
1	80 (8.0%)
2	6 (0.6%)

Query parameter or value	Descriptive statistics
3	1 (0.1%)
SELECT_n_of_non_aggr_func__LOG	
0	946 (95%)
1	52 (5.2%)
2	1 (0.1%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__DOW	
0	919 (92%)
1	71 (7.1%)
2	9 (0.9%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__RTRIM	
0	763 (76%)
1	204 (20%)
2	31 (3.1%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	780 (78%)
1	191 (19%)
2	24 (2.4%)
3	5 (0.5%)
SELECT_n_of_non_aggr_func__LOWER	
0	798 (80%)
1	176 (18%)
2	21 (2.1%)
3	5 (0.5%)
SELECT_n_of_non_aggr_func__UPPER	
0	763 (76%)
1	209 (21%)
2	24 (2.4%)
3	4 (0.4%)
SELECT_n_of_non_aggr_func__MONTH	
0	887 (89%)
1	95 (9.5%)
2	15 (1.5%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func__YEAR	
0	899 (90%)
1	80 (8.0%)
2	16 (1.6%)
3	4 (0.4%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__DAY	
0	943 (94%)
1	52 (5.2%)

Query parameter or value	Descriptive statistics
2	5 (0.5%)
SELECT_n_of_non_aggr_func__ROUND	
0	968 (97%)
1	31 (3.1%)
2	1 (0.1%)
SELECT_n_of_non_aggr_func__TRUNC	
0	966 (97%)
1	31 (3.1%)
2	2 (0.2%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__FLOOR	
0	963 (96%)
1	37 (3.7%)
SELECT_n_of_all_non_aggr_func	[0.00, 1.00, 18.00] 2.00 / 2.31
SELECT_n_of_aggr_func__COUNT_DISTINCT	
0	755 (76%)
1	121 (12%)
2	70 (7.0%)
3	32 (3.2%)
4	13 (1.3%)
5	5 (0.5%)
6	2 (0.2%)
7	2 (0.2%)
SELECT_n_of_aggr_func__COUNT	
0	776 (78%)
1	107 (11%)
2	61 (6.1%)
3	34 (3.4%)
4	14 (1.4%)
5	4 (0.4%)
6	4 (0.4%)
SELECT_n_of_aggr_func__MAX	
0	781 (78%)
1	109 (11%)
2	53 (5.3%)
3	30 (3.0%)
4	18 (1.8%)
5	5 (0.5%)
6	4 (0.4%)
SELECT_n_of_aggr_func__AVG	
0	955 (96%)
1	40 (4.0%)
2	5 (0.5%)
SELECT_n_of_aggr_func__MIN	
0	759 (76%)
1	110 (11%)

Query parameter or value	Descriptive statistics
2	67 (6.7%)
3	37 (3.7%)
4	14 (1.4%)
5	7 (0.7%)
6	5 (0.5%)
7	1 (0.1%)
SELECT_n_of_aggr_func__SUM	
0	948 (95%)
1	44 (4.4%)
2	8 (0.8%)
SELECT_n_of_all_aggr_func	
0	173 (17%)
1	293 (29%)
2	233 (23%)
3	153 (15%)
4	82 (8.2%)
5	40 (4.0%)
6	18 (1.8%)
7	8 (0.8%)
FROM_n_of_join_paths	
1	1,000 (100%)
FROM_n_of_joins__RIGHT	
0	465 (47%)
1	345 (35%)
2	146 (15%)
3	42 (4.2%)
4	2 (0.2%)
FROM_n_of_joins__INNER	
0	489 (49%)
1	335 (34%)
2	139 (14%)
3	33 (3.3%)
4	4 (0.4%)
FROM_n_of_processed_rows	[5, 750,000,000, 3,825,028,272] 1,512,480,576 / 1,608,387,537
WHERE_n_of_predicates	
0	240 (24%)
1	211 (21%)
2	212 (21%)
3	189 (19%)
4	88 (8.8%)
5	44 (4.4%)
6	16 (1.6%)
WHERE_n_of_attris_of_type__integer	
0	563 (56%)
1	312 (31%)

Query parameter or value	Descriptive statistics
2	111 (11%)
3	13 (1.3%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type__character_varying	
0	527 (53%)
1	318 (32%)
2	131 (13%)
3	20 (2.0%)
4	4 (0.4%)
WHERE_n_of_attribs_of_type__numeric	
0	734 (73%)
1	224 (22%)
2	33 (3.3%)
3	9 (0.9%)
WHERE_n_of_attribs_of_type__character	
0	827 (83%)
1	154 (15%)
2	19 (1.9%)
WHERE_n_of_attribs_of_type__date	
0	886 (89%)
1	101 (10%)
2	12 (1.2%)
3	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	721 (72%)
1	251 (25%)
2	26 (2.6%)
3	2 (0.2%)
WHERE_n_of_connect_OR	
0	574 (57%)
1	234 (23%)
2	122 (12%)
3	50 (5.0%)
4	16 (1.6%)
5	4 (0.4%)
WHERE_n_of_operators__greater_or_less	
0	521 (52%)
1	330 (33%)
2	112 (11%)
3	29 (2.9%)
4	8 (0.8%)
WHERE_n_of_operators__in	
0	720 (72%)
1	223 (22%)
2	48 (4.8%)
3	9 (0.9%)

Query parameter or value	Descriptive statistics
WHERE_n_of_operators__between	
0	722 (72%)
1	228 (23%)
2	47 (4.7%)
3	2 (0.2%)
4	1 (0.1%)
WHERE_n_of_operators__like	
0	868 (87%)
1	123 (12%)
2	8 (0.8%)
3	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	976 (98%)
1	24 (2.4%)
WHERE_n_of_non_aggr_func__DAY	
0	976 (98%)
1	24 (2.4%)
WHERE_n_of_non_aggr_func__TRUNC	
0	983 (98%)
1	17 (1.7%)
WHERE_n_of_non_aggr_func__DOW	
0	984 (98%)
1	16 (1.6%)
WHERE_n_of_non_aggr_func__SQRT	
0	942 (94%)
1	56 (5.6%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__FLOOR	
0	981 (98%)
1	18 (1.8%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__ABS	
0	981 (98%)
1	19 (1.9%)
WHERE_n_of_non_aggr_func__LOG	
0	965 (97%)
1	34 (3.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__YEAR	
0	985 (99%)
1	14 (1.4%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	

Query parameter or value	Descriptive statistics
<i>0</i>	793 (79%)
<i>1</i>	169 (17%)
<i>2</i>	33 (3.3%)
<i>3</i>	5 (0.5%)
GROUP_BY_n_of_columns	
<i>0</i>	240 (24%)
<i>1</i>	246 (25%)
<i>2</i>	245 (25%)
<i>3</i>	150 (15%)
<i>4</i>	70 (7.0%)
<i>5</i>	41 (4.1%)
<i>6</i>	7 (0.7%)
<i>7</i>	1 (0.1%)
HAVING_n_of_main_predicates	
<i>0</i>	414 (41%)
<i>1</i>	194 (19%)
<i>2</i>	204 (20%)
<i>3</i>	188 (19%)
ORDER_BY_n_of_columns	
<i>0</i>	295 (30%)
<i>1</i>	438 (44%)
<i>2</i>	230 (23%)
<i>3</i>	31 (3.1%)
<i>4</i>	4 (0.4%)
<i>5</i>	2 (0.2%)
limit	[5, 490, 1,000] 489 / 291
offset	[0, 0, 999] 243 / 315