

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

Scenario: 1-1jp__ng

Scale factor: 0.01 GB

The 1000-query pack was generated on: 2023-11-08

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, 43] 10 / 9
SELECT_n_of_non_aggr_func__DAY	
0	913 (91%)
1	69 (6.9%)
2	14 (1.4%)
3	4 (0.4%)
SELECT_n_of_non_aggr_func__MONTH	
0	850 (85%)
1	107 (11%)
2	35 (3.5%)
3	8 (0.8%)
SELECT_n_of_non_aggr_func__ROUND	
0	930 (93%)
1	64 (6.4%)
2	6 (0.6%)

Query parameter or value	Descriptive statistics
SELECT__n_of__non__aggr__func___SQRT	
0	821 (82%)
1	137 (14%)
2	34 (3.4%)
3	7 (0.7%)
4	1 (0.1%)
SELECT__n_of__non__aggr__func___SUBSTR	
0	810 (81%)
1	146 (15%)
2	34 (3.4%)
3	8 (0.8%)
4	1 (0.1%)
6	1 (0.1%)
SELECT__n_of__non__aggr__func___UPPER	
0	789 (79%)
1	164 (16%)
2	37 (3.7%)
3	6 (0.6%)
4	3 (0.3%)
5	1 (0.1%)
SELECT__n_of__non__aggr__func___YEAR	
0	883 (88%)
1	88 (8.8%)
2	24 (2.4%)
3	5 (0.5%)
SELECT__n_of__non__aggr__func___FLOOR	
0	940 (94%)
1	55 (5.5%)
2	4 (0.4%)
3	1 (0.1%)
SELECT__n_of__non__aggr__func___LOWER	
0	791 (79%)
1	172 (17%)
2	33 (3.3%)
3	4 (0.4%)
SELECT__n_of__non__aggr__func___LTRIM	
0	778 (78%)
1	175 (18%)
2	36 (3.6%)
3	7 (0.7%)
4	3 (0.3%)
5	1 (0.1%)
SELECT__n_of__non__aggr__func___RTRIM	
0	794 (79%)
1	146 (15%)
2	47 (4.7%)

Query parameter or value	Descriptive statistics
3	11 (1.1%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func___TRUNC	
0	955 (96%)
1	42 (4.2%)
2	3 (0.3%)
SELECT_n_of_non_aggr_func___ABS	
0	935 (94%)
1	62 (6.2%)
2	2 (0.2%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___DOW	
0	887 (89%)
1	93 (9.3%)
2	17 (1.7%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func___LOG	
0	900 (90%)
1	86 (8.6%)
2	10 (1.0%)
3	4 (0.4%)
SELECT_n_of_all_non_aggr_func	[0.0, 1.0, 22.0] 2.5 / 3.7
SELECT_n_of_aggr_func___MAX	
0	924 (92%)
1	34 (3.4%)
2	20 (2.0%)
3	15 (1.5%)
4	4 (0.4%)
5	3 (0.3%)
SELECT_n_of_aggr_func___COUNT	
0	927 (93%)
1	26 (2.6%)
2	22 (2.2%)
3	19 (1.9%)
4	4 (0.4%)
6	2 (0.2%)
SELECT_n_of_aggr_func___MIN	
0	937 (94%)
1	28 (2.8%)
2	27 (2.7%)
3	3 (0.3%)
4	3 (0.3%)
5	2 (0.2%)
SELECT_n_of_aggr_func___COUNT_DISTINCT	
0	933 (93%)
1	36 (3.6%)

Query parameter or value	Descriptive statistics
2	14 (1.4%)
3	13 (1.3%)
4	2 (0.2%)
5	2 (0.2%)
SELECT_n_of_aggr_func__SUM	
0	986 (99%)
1	9 (0.9%)
2	5 (0.5%)
SELECT_n_of_aggr_func__AVG	
0	989 (99%)
1	10 (1.0%)
2	1 (0.1%)
SELECT_n_of_all_aggr_func	
0	750 (75%)
1	84 (8.4%)
2	73 (7.3%)
3	52 (5.2%)
4	28 (2.8%)
5	7 (0.7%)
6	5 (0.5%)
7	1 (0.1%)
FROM_n_of_join_paths	
1	1,000 (100%)
FROM_n_of_joins__INNER	
0	480 (48%)
1	348 (35%)
2	134 (13%)
3	35 (3.5%)
4	3 (0.3%)
FROM_n_of_joins__RIGHT	
0	492 (49%)
1	342 (34%)
2	126 (13%)
3	34 (3.4%)
4	6 (0.6%)
FROM_n_of_processed_rows	[5, 10,000, 76,705] 26,723 / 30,626
WHERE_n_of_predicates	
0	215 (22%)
1	245 (25%)
2	226 (23%)
3	158 (16%)
4	98 (9.8%)
5	41 (4.1%)
6	17 (1.7%)
WHERE_n_of_attribs_of_type__integer	
0	543 (54%)

Query parameter or value	Descriptive statistics
1	329 (33%)
2	110 (11%)
3	16 (1.6%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type__date	
0	891 (89%)
1	98 (9.8%)
2	10 (1.0%)
3	1 (0.1%)
WHERE_n_of_attribs_of_type__numeric	
0	736 (74%)
1	213 (21%)
2	47 (4.7%)
3	4 (0.4%)
WHERE_n_of_attribs_of_type__character	
0	834 (83%)
1	131 (13%)
2	32 (3.2%)
3	3 (0.3%)
WHERE_n_of_attribs_of_type__character_varying	
0	541 (54%)
1	325 (33%)
2	108 (11%)
3	24 (2.4%)
4	2 (0.2%)
WHERE_n_of_pkey_attribs	
0	690 (69%)
1	257 (26%)
2	51 (5.1%)
3	2 (0.2%)
WHERE_n_of_connect_OR	
0	590 (59%)
1	224 (22%)
2	113 (11%)
3	54 (5.4%)
4	11 (1.1%)
5	8 (0.8%)
WHERE_n_of_operators__greater_or_less	
0	528 (53%)
1	314 (31%)
2	123 (12%)
3	29 (2.9%)
4	6 (0.6%)
WHERE_n_of_operators__between	
0	699 (70%)
1	247 (25%)

Query parameter or value	Descriptive statistics
2	48 (4.8%)
3	6 (0.6%)
WHERE_n_of_operators_like	
0	880 (88%)
1	114 (11%)
2	6 (0.6%)
WHERE_n_of_operators_in	
0	692 (69%)
1	255 (26%)
2	48 (4.8%)
3	5 (0.5%)
WHERE_n_of_non_aggr_func_ABS	
0	978 (98%)
1	22 (2.2%)
WHERE_n_of_non_aggr_func_MONTH	
0	983 (98%)
1	17 (1.7%)
WHERE_n_of_non_aggr_func_LOG	
0	965 (97%)
1	34 (3.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_SQRT	
0	936 (94%)
1	61 (6.1%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func_DOW	
0	984 (98%)
1	16 (1.6%)
WHERE_n_of_non_aggr_func_TRUNC	
0	982 (98%)
1	18 (1.8%)
WHERE_n_of_non_aggr_func_ROUND	
0	976 (98%)
1	24 (2.4%)
WHERE_n_of_non_aggr_func_DAY	
0	983 (98%)
1	16 (1.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_YEAR	
0	983 (98%)
1	17 (1.7%)
WHERE_n_of_non_aggr_func_FLOOR	
0	977 (98%)
1	23 (2.3%)
WHERE_n_of_all_non_aggr_func	
0	785 (79%)

Query parameter or value	Descriptive statistics
<i>1</i>	174 (17%)
<i>2</i>	39 (3.9%)
<i>3</i>	2 (0.2%)
GROUP_BY_n_of_columns	
<i>0</i>	1,000 (100%)
HAVING_n_of_main_predicates	
<i>0</i>	1,000 (100%)
ORDER_BY_n_of_columns	
<i>0</i>	439 (44%)
<i>1</i>	287 (29%)
<i>2</i>	161 (16%)
<i>3</i>	80 (8.0%)
<i>4</i>	29 (2.9%)
<i>5</i>	4 (0.4%)
limit	[4, 513, 1,000] 500 / 289
offset	[0, 0, 1,000] 252 / 324