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

Scenario: 1-1jp__ng

Scale factor: 1 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] 9 / 9
SELECT_n_of_non_aggr_func___LTRIM	
0	782 (78%)
1	176 (18%)
2	31 (3.1%)
3	7 (0.7%)
4	3 (0.3%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___UPPER	
0	808 (81%)
1	148 (15%)
2	33 (3.3%)
3	7 (0.7%)
4	4 (0.4%)
SELECT_n_of_non_aggr_func___DAY	

Query parameter or value	Descriptive statistics
0	916 (92%)
1	66 (6.6%)
2	16 (1.6%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___LOWER	
0	802 (80%)
1	154 (15%)
2	33 (3.3%)
3	10 (1.0%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___MONTH	
0	839 (84%)
1	112 (11%)
2	38 (3.8%)
3	10 (1.0%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___ROUND	
0	940 (94%)
1	57 (5.7%)
2	3 (0.3%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	761 (76%)
1	196 (20%)
2	36 (3.6%)
3	5 (0.5%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___YEAR	
0	865 (87%)
1	105 (11%)
2	24 (2.4%)
3	5 (0.5%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___SQRT	
0	841 (84%)
1	132 (13%)
2	18 (1.8%)
3	8 (0.8%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___FLOOR	
0	945 (95%)
1	51 (5.1%)
2	2 (0.2%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___LOG	
0	901 (90%)

Query parameter or value	Descriptive statistics
1	86 (8.6%)
2	13 (1.3%)
SELECT_n_of_non_aggr_func___RTRIM	
0	791 (79%)
1	155 (16%)
2	43 (4.3%)
3	8 (0.8%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___ABS	
0	934 (93%)
1	58 (5.8%)
2	8 (0.8%)
SELECT_n_of_non_aggr_func___DOW	
0	893 (89%)
1	85 (8.5%)
2	20 (2.0%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___TRUNC	
0	938 (94%)
1	58 (5.8%)
2	4 (0.4%)
SELECT_n_of_all_non_aggr_func	[0.0, 1.0, 20.0] 2.6 / 3.6
SELECT_n_of_aggr_func___MIN	
0	928 (93%)
1	36 (3.6%)
2	19 (1.9%)
3	10 (1.0%)
4	5 (0.5%)
5	2 (0.2%)
SELECT_n_of_aggr_func___COUNT_DISTINCT	
0	920 (92%)
1	42 (4.2%)
2	19 (1.9%)
3	8 (0.8%)
4	7 (0.7%)
5	3 (0.3%)
6	1 (0.1%)
SELECT_n_of_aggr_func___MAX	
0	930 (93%)
1	29 (2.9%)
2	18 (1.8%)
3	15 (1.5%)
4	5 (0.5%)
5	1 (0.1%)
6	1 (0.1%)

Query parameter or value	Descriptive statistics
7	1 (0.1%)
SELECT_n_of_aggr_func___COUNT	
0	918 (92%)
1	31 (3.1%)
2	22 (2.2%)
3	16 (1.6%)
4	9 (0.9%)
5	3 (0.3%)
6	1 (0.1%)
SELECT_n_of_aggr_func___AVG	
0	985 (99%)
1	12 (1.2%)
2	3 (0.3%)
SELECT_n_of_aggr_func___SUM	
0	988 (99%)
1	11 (1.1%)
3	1 (0.1%)
SELECT_n_of_all_aggr_func	
0	731 (73%)
1	90 (9.0%)
2	68 (6.8%)
3	60 (6.0%)
4	22 (2.2%)
5	23 (2.3%)
6	4 (0.4%)
7	2 (0.2%)
FROM_n_of_join_paths	
1	1,000 (100%)
FROM_n_of_joins___RIGHT	
0	489 (49%)
1	319 (32%)
2	158 (16%)
3	28 (2.8%)
4	6 (0.6%)
FROM_n_of_joins___INNER	
0	502 (50%)
1	330 (33%)
2	133 (13%)
3	31 (3.1%)
4	4 (0.4%)
FROM_n_of_processed_rows	[5, 1,000,000, 7,650,621] 2,746,887 / 3,102,421
WHERE_n_of_predicates	
0	257 (26%)
1	231 (23%)
2	212 (21%)
3	167 (17%)

Query parameter or value	Descriptive statistics
4	79 (7.9%)
5	43 (4.3%)
6	11 (1.1%)
WHERE_n_of_attribs_of_type___character_varying	
0	567 (57%)
1	321 (32%)
2	96 (9.6%)
3	16 (1.6%)
WHERE_n_of_attribs_of_type___integer	
0	561 (56%)
1	327 (33%)
2	96 (9.6%)
3	13 (1.3%)
4	3 (0.3%)
WHERE_n_of_attribs_of_type___character	
0	816 (82%)
1	157 (16%)
2	25 (2.5%)
3	1 (0.1%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type___numeric	
0	753 (75%)
1	203 (20%)
2	41 (4.1%)
3	3 (0.3%)
WHERE_n_of_attribs_of_type___date	
0	900 (90%)
1	87 (8.7%)
2	12 (1.2%)
3	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	717 (72%)
1	248 (25%)
2	34 (3.4%)
3	1 (0.1%)
WHERE_n_of_connect_OR	
0	626 (63%)
1	204 (20%)
2	105 (11%)
3	44 (4.4%)
4	20 (2.0%)
5	1 (0.1%)
WHERE_n_of_operators___greater_or_less	
0	547 (55%)
1	288 (29%)
2	124 (12%)

Query parameter or value	Descriptive statistics
3	37 (3.7%)
4	4 (0.4%)
WHERE_n_of_operators___between	
0	739 (74%)
1	216 (22%)
2	39 (3.9%)
3	5 (0.5%)
4	1 (0.1%)
WHERE_n_of_operators___in	
0	729 (73%)
1	228 (23%)
2	36 (3.6%)
3	6 (0.6%)
4	1 (0.1%)
WHERE_n_of_operators___like	
0	884 (88%)
1	110 (11%)
2	6 (0.6%)
WHERE_n_of_non_aggr_func___FLOOR	
0	986 (99%)
1	14 (1.4%)
WHERE_n_of_non_aggr_func___TRUNC	
0	984 (98%)
1	15 (1.5%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___DAY	
0	987 (99%)
1	13 (1.3%)
WHERE_n_of_non_aggr_func___SQRT	
0	933 (93%)
1	66 (6.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___LOG	
0	954 (95%)
1	46 (4.6%)
WHERE_n_of_non_aggr_func___YEAR	
0	984 (98%)
1	14 (1.4%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___ROUND	
0	978 (98%)
1	22 (2.2%)
WHERE_n_of_non_aggr_func___ABS	
0	989 (99%)
1	11 (1.1%)
WHERE_n_of_non_aggr_func___DOW	

Query parameter or value	Descriptive statistics
0	985 (99%)
1	15 (1.5%)
WHERE_n_of_non_aggr_func___MONTH	
0	980 (98%)
1	20 (2.0%)
WHERE_n_of_all_non_aggr_func	
0	790 (79%)
1	181 (18%)
2	24 (2.4%)
3	5 (0.5%)
GROUP_BY_n_of_columns	
0	1,000 (100%)
HAVING_n_of_main_predicates	
0	1,000 (100%)
ORDER_BY_n_of_columns	
0	465 (47%)
1	273 (27%)
2	157 (16%)
3	66 (6.6%)
4	27 (2.7%)
5	11 (1.1%)
6	1 (0.1%)
limit	[2, 517, 1,000] 507 / 293
offset	[0, 42, 998] 270 / 332