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

Scenario: 1-1jp__ng

Scale factor: 200 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, 5, 40] 9 / 9
SELECT_n_of_non_aggr_func___LOWER	
0	783 (78%)
1	170 (17%)
2	37 (3.7%)
3	8 (0.8%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func___UPPER	
0	802 (80%)
1	150 (15%)
2	38 (3.8%)
3	10 (1.0%)
SELECT_n_of_non_aggr_func___ROUND	
0	952 (95%)
1	46 (4.6%)

Query parameter or value	Descriptive statistics
2	2 (0.2%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	781 (78%)
1	172 (17%)
2	37 (3.7%)
3	9 (0.9%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___ABS	
0	942 (94%)
1	51 (5.1%)
2	7 (0.7%)
SELECT_n_of_non_aggr_func___DOW	
0	898 (90%)
1	82 (8.2%)
2	18 (1.8%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___LTRIM	
0	784 (78%)
1	170 (17%)
2	41 (4.1%)
3	4 (0.4%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___RTRIM	
0	777 (78%)
1	176 (18%)
2	34 (3.4%)
3	8 (0.8%)
4	4 (0.4%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___SQRT	
0	836 (84%)
1	130 (13%)
2	27 (2.7%)
3	7 (0.7%)
SELECT_n_of_non_aggr_func___LOG	
0	914 (91%)
1	72 (7.2%)
2	12 (1.2%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___MONTH	
0	861 (86%)
1	106 (11%)
2	20 (2.0%)
3	13 (1.3%)
SELECT_n_of_non_aggr_func___YEAR	
0	885 (89%)

Query parameter or value	Descriptive statistics
1	90 (9.0%)
2	18 (1.8%)
3	6 (0.6%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___DAY	
0	919 (92%)
1	69 (6.9%)
2	11 (1.1%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___TRUNC	
0	939 (94%)
1	54 (5.4%)
2	7 (0.7%)
SELECT_n_of_non_aggr_func___FLOOR	
0	948 (95%)
1	45 (4.5%)
2	7 (0.7%)
SELECT_n_of_all_non_aggr_func	[0.00, 1.00, 20.00] 2.47 / 3.47
SELECT_n_of_aggr_func___COUNT_DISTINCT	
0	911 (91%)
1	42 (4.2%)
2	22 (2.2%)
3	16 (1.6%)
4	7 (0.7%)
5	2 (0.2%)
SELECT_n_of_aggr_func___MIN	
0	939 (94%)
1	37 (3.7%)
2	15 (1.5%)
3	4 (0.4%)
4	2 (0.2%)
5	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_aggr_func___COUNT	
0	920 (92%)
1	32 (3.2%)
2	26 (2.6%)
3	9 (0.9%)
4	10 (1.0%)
5	1 (0.1%)
6	2 (0.2%)
SELECT_n_of_aggr_func___MAX	
0	940 (94%)
1	27 (2.7%)
2	16 (1.6%)
3	10 (1.0%)

Query parameter or value	Descriptive statistics
4	6 (0.6%)
5	1 (0.1%)
SELECT_n_of_aggr_func___AVG	
0	985 (99%)
1	13 (1.3%)
2	2 (0.2%)
SELECT_n_of_aggr_func___SUM	
0	987 (99%)
1	13 (1.3%)
SELECT_n_of_all_aggr_func	
0	742 (74%)
1	93 (9.3%)
2	74 (7.4%)
3	41 (4.1%)
4	31 (3.1%)
5	11 (1.1%)
6	7 (0.7%)
7	1 (0.1%)
FROM_n_of_join_paths	
1	1,000 (100%)
FROM_n_of_joins___INNER	
0	526 (53%)
1	315 (32%)
2	127 (13%)
3	28 (2.8%)
4	4 (0.4%)
FROM_n_of_joins___RIGHT	
0	512 (51%)
1	312 (31%)
2	140 (14%)
3	30 (3.0%)
4	6 (0.6%)
FROM_n_of_processed_rows	[5, 200,000,000, 1,530,018,464] 545,216,114 / 617,980,396
WHERE_n_of_predicates	
0	269 (27%)
1	226 (23%)
2	199 (20%)
3	166 (17%)
4	97 (9.7%)
5	35 (3.5%)
6	8 (0.8%)
WHERE_n_of_attribs_of_type___numeric	
0	738 (74%)
1	217 (22%)
2	41 (4.1%)

Query parameter or value	Descriptive statistics
3	4 (0.4%)
WHERE_n_of_attribs_of_type___character_varying	
0	557 (56%)
1	335 (34%)
2	88 (8.8%)
3	19 (1.9%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type___integer	
0	582 (58%)
1	312 (31%)
2	88 (8.8%)
3	18 (1.8%)
WHERE_n_of_attribs_of_type___character	
0	828 (83%)
1	145 (15%)
2	25 (2.5%)
3	2 (0.2%)
WHERE_n_of_attribs_of_type___date	
0	897 (90%)
1	99 (9.9%)
2	4 (0.4%)
WHERE_n_of_pkey_attribs	
0	718 (72%)
1	253 (25%)
2	25 (2.5%)
3	4 (0.4%)
WHERE_n_of_connect_OR	
0	629 (63%)
1	208 (21%)
2	97 (9.7%)
3	52 (5.2%)
4	13 (1.3%)
5	1 (0.1%)
WHERE_n_of_operators___greater_or_less	
0	564 (56%)
1	289 (29%)
2	116 (12%)
3	29 (2.9%)
4	2 (0.2%)
WHERE_n_of_operators___in	
0	712 (71%)
1	241 (24%)
2	41 (4.1%)
3	6 (0.6%)
WHERE_n_of_operators___between	
0	751 (75%)

Query parameter or value	Descriptive statistics
1	209 (21%)
2	36 (3.6%)
3	4 (0.4%)
WHERE_n_of_operators___like	
0	897 (90%)
1	98 (9.8%)
2	5 (0.5%)
WHERE_n_of_non_aggr_func___ROUND	
0	979 (98%)
1	21 (2.1%)
WHERE_n_of_non_aggr_func___ABS	
0	987 (99%)
1	13 (1.3%)
WHERE_n_of_non_aggr_func___LOG	
0	964 (96%)
1	36 (3.6%)
WHERE_n_of_non_aggr_func___SQRT	
0	926 (93%)
1	71 (7.1%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func___TRUNC	
0	982 (98%)
1	18 (1.8%)
WHERE_n_of_non_aggr_func___DAY	
0	976 (98%)
1	24 (2.4%)
WHERE_n_of_non_aggr_func___FLOOR	
0	978 (98%)
1	22 (2.2%)
WHERE_n_of_non_aggr_func___YEAR	
0	985 (99%)
1	15 (1.5%)
WHERE_n_of_non_aggr_func___DOW	
0	991 (99%)
1	9 (0.9%)
WHERE_n_of_non_aggr_func___MONTH	
0	987 (99%)
1	13 (1.3%)
WHERE_n_of_all_non_aggr_func	
0	788 (79%)
1	177 (18%)
2	34 (3.4%)
3	1 (0.1%)
GROUP_BY_n_of_columns	
0	1,000 (100%)
HAVING_n_of_main_predicates	

Query parameter or value	Descriptive statistics
0	1,000 (100%)
ORDER_BY_n_of_columns	
0	448 (45%)
1	289 (29%)
2	154 (15%)
3	68 (6.8%)
4	30 (3.0%)
5	10 (1.0%)
6	1 (0.1%)
limit	[5, 529, 1,000] 525 / 282
offset	[0, 29, 997] 262 / 330