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

Scenario: 1-1jp__g__h0ss_h0nss

Scale factor: 0.01 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] 5.5 / 5.1
SELECT_n_of_non_aggr_func___RTRIM	
0	787 (79%)
1	183 (18%)
2	28 (2.8%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___UPPER	
0	796 (80%)
1	184 (18%)
2	15 (1.5%)
3	5 (0.5%)
SELECT_n_of_non_aggr_func___LTRIM	
0	771 (77%)
1	203 (20%)
2	20 (2.0%)

Query parameter or value	Descriptive statistics
3	6 (0.6%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	778 (78%)
1	192 (19%)
2	28 (2.8%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___SQRT	
0	897 (90%)
1	94 (9.4%)
2	8 (0.8%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___LOWER	
0	768 (77%)
1	191 (19%)
2	36 (3.6%)
3	3 (0.3%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func___LOG	
0	922 (92%)
1	67 (6.7%)
2	9 (0.9%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___ABS	
0	972 (97%)
1	28 (2.8%)
SELECT_n_of_non_aggr_func___DAY	
0	947 (95%)
1	51 (5.1%)
2	2 (0.2%)
SELECT_n_of_non_aggr_func___MONTH	
0	895 (90%)
1	93 (9.3%)
2	10 (1.0%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___DOW	
0	921 (92%)
1	70 (7.0%)
2	8 (0.8%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___YEAR	
0	910 (91%)
1	79 (7.9%)
2	9 (0.9%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___FLOOR	

Query parameter or value	Descriptive statistics
0	966 (97%)
1	34 (3.4%)
SELECT_n_of_non_aggr_func___ROUND	
0	962 (96%)
1	38 (3.8%)
SELECT_n_of_non_aggr_func___TRUNC	
0	972 (97%)
1	28 (2.8%)
SELECT_n_of_all_non_aggr_func	[0.00, 1.00, 18.00] 1.97 / 2.11
SELECT_n_of_aggr_func___COUNT_DISTINCT	
0	751 (75%)
1	121 (12%)
2	59 (5.9%)
3	41 (4.1%)
4	19 (1.9%)
5	4 (0.4%)
6	3 (0.3%)
7	2 (0.2%)
SELECT_n_of_aggr_func___MAX	
0	757 (76%)
1	130 (13%)
2	54 (5.4%)
3	35 (3.5%)
4	12 (1.2%)
5	5 (0.5%)
6	3 (0.3%)
7	3 (0.3%)
9	1 (0.1%)
SELECT_n_of_aggr_func___COUNT	
0	760 (76%)
1	121 (12%)
2	63 (6.3%)
3	31 (3.1%)
4	15 (1.5%)
5	6 (0.6%)
6	4 (0.4%)
SELECT_n_of_aggr_func___MIN	
0	786 (79%)
1	108 (11%)
2	56 (5.6%)
3	32 (3.2%)
4	11 (1.1%)
5	7 (0.7%)
SELECT_n_of_aggr_func___SUM	
0	954 (95%)
1	38 (3.8%)

Query parameter or value	Descriptive statistics
2	8 (0.8%)
SELECT_n_of_aggr_func___AVG	
0	963 (96%)
1	32 (3.2%)
2	4 (0.4%)
3	1 (0.1%)
SELECT_n_of_all_aggr_func	
0	163 (16%)
1	317 (32%)
2	221 (22%)
3	158 (16%)
4	72 (7.2%)
5	45 (4.5%)
6	16 (1.6%)
7	7 (0.7%)
9	1 (0.1%)
FROM_n_of_join_paths	
1	1,000 (100%)
FROM_n_of_joins___RIGHT	
0	473 (47%)
1	350 (35%)
2	138 (14%)
3	34 (3.4%)
4	5 (0.5%)
FROM_n_of_joins___INNER	
0	486 (49%)
1	336 (34%)
2	131 (13%)
3	43 (4.3%)
4	4 (0.4%)
FROM_n_of_processed_rows	[5, 10,000, 76,705] 28,814 / 31,556
WHERE_n_of_predicates	
0	228 (23%)
1	240 (24%)
2	185 (19%)
3	177 (18%)
4	99 (9.9%)
5	56 (5.6%)
6	15 (1.5%)
WHERE_n_of_attribs_of_type___integer	
0	563 (56%)
1	310 (31%)
2	100 (10%)
3	25 (2.5%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type___character_varying	

Query parameter or value	Descriptive statistics
0	538 (54%)
1	317 (32%)
2	120 (12%)
3	23 (2.3%)
4	1 (0.1%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type___numeric	
0	724 (72%)
1	237 (24%)
2	37 (3.7%)
3	2 (0.2%)
WHERE_n_of_attribs_of_type___character	
0	815 (82%)
1	161 (16%)
2	22 (2.2%)
3	2 (0.2%)
WHERE_n_of_attribs_of_type___date	
0	872 (87%)
1	107 (11%)
2	19 (1.9%)
3	2 (0.2%)
WHERE_n_of_pkey_attribs	
0	719 (72%)
1	240 (24%)
2	38 (3.8%)
3	2 (0.2%)
4	1 (0.1%)
WHERE_n_of_connect_OR	
0	591 (59%)
1	219 (22%)
2	114 (11%)
3	54 (5.4%)
4	19 (1.9%)
5	3 (0.3%)
WHERE_n_of_operators___greater_or_less	
0	509 (51%)
1	314 (31%)
2	132 (13%)
3	40 (4.0%)
4	3 (0.3%)
5	2 (0.2%)
WHERE_n_of_operators___in	
0	723 (72%)
1	226 (23%)
2	46 (4.6%)
3	5 (0.5%)

Query parameter or value	Descriptive statistics
WHERE_n_of_operators___like	
0	888 (89%)
1	105 (11%)
2	7 (0.7%)
WHERE_n_of_operators___between	
0	689 (69%)
1	252 (25%)
2	50 (5.0%)
3	8 (0.8%)
4	1 (0.1%)
WHERE_n_of_non_aggr_func___LOG	
0	957 (96%)
1	42 (4.2%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___TRUNC	
0	982 (98%)
1	17 (1.7%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func___YEAR	
0	972 (97%)
1	28 (2.8%)
WHERE_n_of_non_aggr_func___MONTH	
0	981 (98%)
1	19 (1.9%)
WHERE_n_of_non_aggr_func___SQRT	
0	936 (94%)
1	61 (6.1%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func___DAY	
0	972 (97%)
1	27 (2.7%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___ABS	
0	984 (98%)
1	16 (1.6%)
WHERE_n_of_non_aggr_func___ROUND	
0	981 (98%)
1	19 (1.9%)
WHERE_n_of_non_aggr_func___FLOOR	
0	980 (98%)
1	20 (2.0%)
WHERE_n_of_non_aggr_func___DOW	
0	987 (99%)
1	13 (1.3%)
WHERE_n_of_all_non_aggr_func	
0	764 (76%)

Query parameter or value	Descriptive statistics
1	201 (20%)
2	31 (3.1%)
3	4 (0.4%)
GROUP_BY_n_of_columns	
0	226 (23%)
1	273 (27%)
2	239 (24%)
3	145 (15%)
4	72 (7.2%)
5	41 (4.1%)
6	4 (0.4%)
HAVING_n_of_main_predicates	
0	420 (42%)
1	198 (20%)
2	181 (18%)
3	201 (20%)
ORDER_BY_n_of_columns	
0	288 (29%)
1	503 (50%)
2	180 (18%)
3	23 (2.3%)
4	4 (0.4%)
5	2 (0.2%)
limit	[2, 502, 1,000] 500 / 292
offset	[0, 0, 1,000] 245 / 324