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

Scenario: 1-6jp_rjjp__g__h0ss_h0nss

Scale factor: 5 GB

The 1000-query pack was generated on: 2023-11-12

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, 9, 168] 17 / 26
SELECT_n_of_non_aggr_func___LOG	
0	864 (86%)
1	98 (9.8%)
2	23 (2.3%)
3	8 (0.8%)
4	2 (0.2%)
5	4 (0.4%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___DOW	
0	847 (85%)
1	111 (11%)
2	20 (2.0%)
3	12 (1.2%)
4	4 (0.4%)

Query parameter or value	Descriptive statistics
5	3 (0.3%)
6	2 (0.2%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___LTRIM	[0.00, 0.00, 9.00] 0.63 / 1.13
SELECT_n_of_non_aggr_func___RTRIM	
0	606 (61%)
1	258 (26%)
2	82 (8.2%)
3	27 (2.7%)
4	12 (1.2%)
5	6 (0.6%)
6	5 (0.5%)
7	4 (0.4%)
SELECT_n_of_non_aggr_func___SQRT	
0	796 (80%)
1	145 (15%)
2	30 (3.0%)
3	11 (1.1%)
4	9 (0.9%)
5	4 (0.4%)
6	1 (0.1%)
7	2 (0.2%)
8	2 (0.2%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	608 (61%)
1	275 (28%)
10	2 (0.2%)
2	66 (6.6%)
3	29 (2.9%)
4	8 (0.8%)
5	3 (0.3%)
6	6 (0.6%)
7	3 (0.3%)
SELECT_n_of_non_aggr_func___UPPER	
0	617 (62%)
1	275 (28%)
2	55 (5.5%)
3	23 (2.3%)
4	15 (1.5%)
5	7 (0.7%)
6	2 (0.2%)
7	5 (0.5%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___YEAR	[0.00, 0.00, 10.00] 0.30 / 0.89
SELECT_n_of_non_aggr_func___LOWER	[0.00, 0.00, 9.00] 0.63 / 1.07
SELECT_n_of_non_aggr_func___TRUNC	

Query parameter or value	Descriptive statistics
0	919 (92%)
1	64 (6.4%)
2	14 (1.4%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func___ABS	
0	924 (92%)
1	62 (6.2%)
2	11 (1.1%)
3	2 (0.2%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___DAY	
0	866 (87%)
1	100 (10%)
2	19 (1.9%)
3	7 (0.7%)
4	4 (0.4%)
5	3 (0.3%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___MONTH	
0	802 (80%)
1	121 (12%)
10	1 (0.1%)
2	35 (3.5%)
3	19 (1.9%)
4	9 (0.9%)
5	6 (0.6%)
6	5 (0.5%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func___ROUND	
0	929 (93%)
1	59 (5.9%)
2	9 (0.9%)
3	1 (0.1%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func___FLOOR	
0	921 (92%)
1	65 (6.5%)
2	11 (1.1%)
3	2 (0.2%)
4	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 77.0] 5.1 / 8.5
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 14.00] 0.91 / 2.08
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 11.00] 0.87 / 2.02
SELECT_n_of_aggr_func___AVG	
0	922 (92%)
1	47 (4.7%)

Query parameter or value	Descriptive statistics
2	21 (2.1%)
3	8 (0.8%)
4	2 (0.2%)
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 13.00] 0.91 / 2.14
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 14.00] 0.77 / 1.80
SELECT_n_of_aggr_func___SUM	
0	939 (94%)
1	41 (4.1%)
2	16 (1.6%)
3	1 (0.1%)
4	3 (0.3%)
SELECT_n_of_all_aggr_func	[0.0, 3.0, 15.0] 3.7 / 3.3
FROM_n_of_join_paths	
1	154 (15%)
2	173 (17%)
3	165 (17%)
4	166 (17%)
5	172 (17%)
6	170 (17%)
FROM_n_of_super_joins___FULL	
0	531 (53%)
1	336 (34%)
2	112 (11%)
3	20 (2.0%)
4	1 (0.1%)
FROM_n_of_super_joins___RIGHT	
0	559 (56%)
1	296 (30%)
2	112 (11%)
3	30 (3.0%)
4	2 (0.2%)
5	1 (0.1%)
FROM_n_of_super_joins___LEFT	
0	529 (53%)
1	326 (33%)
2	119 (12%)
3	23 (2.3%)
4	2 (0.2%)
5	1 (0.1%)
FROM_n_of_joins___INNER	[0.00, 3.00, 12.00] 3.29 / 2.49
FROM_n_of_joins___RIGHT	[0.00, 3.00, 13.00] 3.40 / 2.53
FROM_n_of_processed_rows	[5, 46,500,091, 227,999,584] 59,940,429 / 49,179,672
WHERE_n_of_predicates	[0.00, 3.00, 13.00] 3.84 / 2.97
WHERE_n_of_attribs_of_type___character	
0	663 (66%)
1	238 (24%)

Query parameter or value	Descriptive statistics
2	76 (7.6%)
3	19 (1.9%)
4	3 (0.3%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type___character_varying	
0	368 (37%)
1	271 (27%)
2	177 (18%)
3	115 (12%)
4	42 (4.2%)
5	17 (1.7%)
6	10 (1.0%)
WHERE_n_of_attribs_of_type___integer	
0	366 (37%)
1	287 (29%)
2	191 (19%)
3	94 (9.4%)
4	43 (4.3%)
5	12 (1.2%)
6	6 (0.6%)
7	1 (0.1%)
WHERE_n_of_attribs_of_type___numeric	
0	588 (59%)
1	275 (28%)
2	93 (9.3%)
3	36 (3.6%)
4	7 (0.7%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type___date	
0	783 (78%)
1	176 (18%)
2	34 (3.4%)
3	6 (0.6%)
4	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	523 (52%)
1	326 (33%)
2	105 (11%)
3	36 (3.6%)
4	8 (0.8%)
5	1 (0.1%)
6	1 (0.1%)
WHERE_n_of_connect_OR	[0.00, 1.00, 12.00] 2.18 / 2.43
WHERE_n_of_operators___between	
0	518 (52%)
1	303 (30%)

Query parameter or value	Descriptive statistics
2	117 (12%)
3	46 (4.6%)
4	12 (1.2%)
5	4 (0.4%)
WHERE_n_of_operators___greater_or_less	[0.00, 1.00, 9.00] 1.42 / 1.43
WHERE_n_of_operators___like	
0	777 (78%)
1	183 (18%)
2	36 (3.6%)
3	4 (0.4%)
WHERE_n_of_operators___in	
0	542 (54%)
1	298 (30%)
2	114 (11%)
3	33 (3.3%)
4	10 (1.0%)
5	3 (0.3%)
WHERE_n_of_non_aggr_func___FLOOR	
0	960 (96%)
1	40 (4.0%)
WHERE_n_of_non_aggr_func___ROUND	
0	962 (96%)
1	38 (3.8%)
WHERE_n_of_non_aggr_func___SQRT	
0	874 (87%)
1	119 (12%)
2	7 (0.7%)
WHERE_n_of_non_aggr_func___TRUNC	
0	950 (95%)
1	49 (4.9%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___LOG	
0	916 (92%)
1	80 (8.0%)
2	3 (0.3%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func___DOW	
0	965 (97%)
1	35 (3.5%)
WHERE_n_of_non_aggr_func___ABS	
0	956 (96%)
1	42 (4.2%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___MONTH	
0	957 (96%)
1	42 (4.2%)

Query parameter or value	Descriptive statistics
2	1 (0.1%)
WHERE_n_of_non_aggr_func___DAY	
0	968 (97%)
1	28 (2.8%)
2	4 (0.4%)
WHERE_n_of_non_aggr_func___YEAR	
0	969 (97%)
1	30 (3.0%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	607 (61%)
1	271 (27%)
2	97 (9.7%)
3	21 (2.1%)
4	4 (0.4%)
GROUP_BY_n_of_columns	[0.00, 2.00, 13.00] 3.14 / 2.94
HAVING_n_of_main_predicates	
0	450 (45%)
1	179 (18%)
2	193 (19%)
3	178 (18%)
ORDER_BY_n_of_columns	
0	284 (28%)
1	349 (35%)
2	242 (24%)
3	82 (8.2%)
4	25 (2.5%)
5	10 (1.0%)
6	6 (0.6%)
8	2 (0.2%)
limit	[1, 497, 1,000] 491 / 289
offset	[0, 7, 1,000] 256 / 329