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

Scenario: 1-6jp_ujjp__g__h0ss_h0nss

Scale factor: 500 GB

The 1000-query pack was generated on: 2023-11-10

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, 210] 16 / 25
SELECT_n_of_non_aggr_func___DOW	
0	852 (85%)
1	104 (10%)
2	20 (2.0%)
3	10 (1.0%)
4	9 (0.9%)
5	3 (0.3%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func___MONTH	
0	802 (80%)
1	134 (13%)
11	1 (0.1%)
2	37 (3.7%)
3	7 (0.7%)

Query parameter or value	Descriptive statistics
4	10 (1.0%)
5	5 (0.5%)
6	3 (0.3%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___RTRIM	[0.00, 0.00, 14.00] 0.62 / 1.22
SELECT_n_of_non_aggr_func___SQRT	
0	811 (81%)
1	132 (13%)
2	27 (2.7%)
3	14 (1.4%)
4	7 (0.7%)
5	3 (0.3%)
6	4 (0.4%)
7	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___UPPER	[0.00, 0.00, 13.00] 0.64 / 1.19
SELECT_n_of_non_aggr_func___LOG	
0	875 (88%)
1	89 (8.9%)
2	24 (2.4%)
3	9 (0.9%)
4	1 (0.1%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___LTRIM	
0	607 (61%)
1	269 (27%)
2	74 (7.4%)
3	29 (2.9%)
4	8 (0.8%)
5	6 (0.6%)
6	2 (0.2%)
8	4 (0.4%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func___SUBSTR	[0.00, 0.00, 10.00] 0.59 / 1.07
SELECT_n_of_non_aggr_func___ABS	
0	912 (91%)
1	74 (7.4%)
2	9 (0.9%)
3	4 (0.4%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___DAY	
0	878 (88%)
1	95 (9.5%)
2	18 (1.8%)
3	5 (0.5%)
4	3 (0.3%)

Query parameter or value	Descriptive statistics
5	1 (0.1%)
SELECT_n_of_non_aggr_func___LOWER	
0	620 (62%)
1	246 (25%)
2	84 (8.4%)
3	27 (2.7%)
4	10 (1.0%)
5	5 (0.5%)
6	4 (0.4%)
7	2 (0.2%)
9	2 (0.2%)
SELECT_n_of_non_aggr_func___ROUND	
0	911 (91%)
1	69 (6.9%)
2	14 (1.4%)
3	5 (0.5%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___YEAR	
0	831 (83%)
1	110 (11%)
2	36 (3.6%)
3	7 (0.7%)
4	5 (0.5%)
5	6 (0.6%)
6	3 (0.3%)
7	1 (0.1%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func___FLOOR	
0	931 (93%)
1	57 (5.7%)
2	8 (0.8%)
3	1 (0.1%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___TRUNC	
0	920 (92%)
1	62 (6.2%)
2	10 (1.0%)
3	4 (0.4%)
4	2 (0.2%)
5	2 (0.2%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 83.0] 5.0 / 8.4
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 11.00] 0.77 / 1.85
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 13.00] 0.97 / 2.17
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 12.00] 0.85 / 1.98
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 11.00] 0.73 / 1.69

Query parameter or value	Descriptive statistics
SELECT_n_of_aggr_func___SUM	
0	938 (94%)
1	39 (3.9%)
2	14 (1.4%)
3	8 (0.8%)
6	1 (0.1%)
SELECT_n_of_aggr_func___AVG	
0	941 (94%)
1	36 (3.6%)
2	17 (1.7%)
3	5 (0.5%)
4	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.00, 3.00, 15.00] 3.50 / 3.15
FROM_n_of_join_paths	
1	160 (16%)
2	166 (17%)
3	155 (16%)
4	206 (21%)
5	155 (16%)
6	158 (16%)
FROM_n_of_super_joins___RIGHT	
0	530 (53%)
1	341 (34%)
2	101 (10%)
3	25 (2.5%)
4	3 (0.3%)
FROM_n_of_super_joins___LEFT	
0	521 (52%)
1	333 (33%)
2	122 (12%)
3	21 (2.1%)
4	3 (0.3%)
FROM_n_of_super_joins___FULL	
0	561 (56%)
1	303 (30%)
2	114 (11%)
3	21 (2.1%)
4	1 (0.1%)
FROM_n_of_joins___INNER	[0.00, 3.00, 14.00] 3.44 / 2.56
FROM_n_of_joins___RIGHT	[0.00, 3.00, 14.00] 3.28 / 2.45
FROM_n_of_processed_rows	[5, 4,647,528,325, 22,625,169,532] 5,867,036,124 / 4,899,106,877
WHERE_n_of_predicates	[0.00, 3.00, 14.00] 3.75 / 3.00
WHERE_n_of_attribs_of_type___character_varying	
0	354 (35%)
1	315 (32%)

Query parameter or value	Descriptive statistics
2	171 (17%)
3	97 (9.7%)
4	41 (4.1%)
5	16 (1.6%)
6	3 (0.3%)
7	3 (0.3%)
WHERE_n_of_attribs_of_type___integer	[0.00, 1.00, 9.00] 1.18 / 1.30
WHERE_n_of_attribs_of_type___character	
0	682 (68%)
1	237 (24%)
2	61 (6.1%)
3	17 (1.7%)
4	3 (0.3%)
WHERE_n_of_attribs_of_type___numeric	
0	582 (58%)
1	262 (26%)
2	110 (11%)
3	38 (3.8%)
4	8 (0.8%)
WHERE_n_of_attribs_of_type___date	
0	773 (77%)
1	179 (18%)
2	35 (3.5%)
3	12 (1.2%)
4	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	544 (54%)
1	287 (29%)
2	121 (12%)
3	36 (3.6%)
4	6 (0.6%)
5	5 (0.5%)
6	1 (0.1%)
WHERE_n_of_connect_OR	[0.00, 1.00, 11.00] 2.16 / 2.44
WHERE_n_of_operators___greater_or_less	
0	337 (34%)
1	264 (26%)
10	1 (0.1%)
2	205 (21%)
3	102 (10%)
4	53 (5.3%)
5	29 (2.9%)
6	7 (0.7%)
7	2 (0.2%)
WHERE_n_of_operators___between	
0	540 (54%)

Query parameter or value	Descriptive statistics
1	281 (28%)
2	122 (12%)
3	39 (3.9%)
4	15 (1.5%)
5	3 (0.3%)
WHERE_n_of_operators___in	
0	558 (56%)
1	276 (28%)
2	111 (11%)
3	40 (4.0%)
4	12 (1.2%)
5	3 (0.3%)
WHERE_n_of_operators___like	
0	805 (81%)
1	165 (17%)
2	24 (2.4%)
3	6 (0.6%)
WHERE_n_of_non_aggr_func___ABS	
0	959 (96%)
1	37 (3.7%)
2	4 (0.4%)
WHERE_n_of_non_aggr_func___SQRT	
0	886 (89%)
1	101 (10%)
2	12 (1.2%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func___FLOOR	
0	968 (97%)
1	31 (3.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___MONTH	
0	958 (96%)
1	41 (4.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___TRUNC	
0	956 (96%)
1	42 (4.2%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___LOG	
0	933 (93%)
1	62 (6.2%)
2	5 (0.5%)
WHERE_n_of_non_aggr_func___YEAR	
0	954 (95%)
1	46 (4.6%)
WHERE_n_of_non_aggr_func___DAY	

Query parameter or value	Descriptive statistics
0	968 (97%)
1	31 (3.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___DOW	
0	949 (95%)
1	50 (5.0%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___ROUND	
0	962 (96%)
1	37 (3.7%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	634 (63%)
1	241 (24%)
2	87 (8.7%)
3	30 (3.0%)
4	8 (0.8%)
GROUP_BY_n_of_columns	[0.00, 3.00, 13.00] 3.08 / 2.91
HAVING_n_of_main_predicates	
0	451 (45%)
1	176 (18%)
2	189 (19%)
3	184 (18%)
ORDER_BY_n_of_columns	
0	297 (30%)
1	323 (32%)
2	239 (24%)
3	98 (9.8%)
4	20 (2.0%)
5	15 (1.5%)
6	4 (0.4%)
7	4 (0.4%)
limit	[2, 512, 1,000] 499 / 292
offset	[0, 24, 999] 261 / 330