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

Scenario: 1-3jp_rjjp__g__h0ss_h0nss

Scale factor: 2 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, 6, 96] 10 / 12
SELECT_n_of_non_aggr_func___DOW	
0	887 (89%)
1	87 (8.7%)
2	16 (1.6%)
3	7 (0.7%)
4	3 (0.3%)
SELECT_n_of_non_aggr_func___FLOOR	
0	953 (95%)
1	43 (4.3%)
2	3 (0.3%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___RTRIM	
0	707 (71%)
1	230 (23%)

Query parameter or value	Descriptive statistics
2	47 (4.7%)
3	11 (1.1%)
4	2 (0.2%)
5	2 (0.2%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	684 (68%)
1	249 (25%)
2	45 (4.5%)
3	17 (1.7%)
4	4 (0.4%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___YEAR	
0	861 (86%)
1	108 (11%)
2	19 (1.9%)
3	7 (0.7%)
4	2 (0.2%)
5	1 (0.1%)
6	2 (0.2%)
SELECT_n_of_non_aggr_func___LOWER	
0	708 (71%)
1	228 (23%)
2	46 (4.6%)
3	13 (1.3%)
4	3 (0.3%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___UPPER	
0	693 (69%)
1	225 (23%)
2	59 (5.9%)
3	19 (1.9%)
4	3 (0.3%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___LTRIM	
0	704 (70%)
1	230 (23%)
2	50 (5.0%)
3	10 (1.0%)
4	4 (0.4%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___ROUND	
0	950 (95%)
1	46 (4.6%)
2	4 (0.4%)
SELECT_n_of_non_aggr_func___TRUNC	

Query parameter or value	Descriptive statistics
0	954 (95%)
1	38 (3.8%)
2	7 (0.7%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___LOG	
0	900 (90%)
1	81 (8.1%)
2	13 (1.3%)
3	6 (0.6%)
SELECT_n_of_non_aggr_func___MONTH	
0	843 (84%)
1	121 (12%)
2	24 (2.4%)
3	7 (0.7%)
5	3 (0.3%)
6	2 (0.2%)
SELECT_n_of_non_aggr_func___ABS	
0	941 (94%)
1	57 (5.7%)
2	2 (0.2%)
SELECT_n_of_non_aggr_func___DAY	
0	914 (91%)
1	72 (7.2%)
2	11 (1.1%)
3	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___SQRT	
0	862 (86%)
1	106 (11%)
2	20 (2.0%)
3	9 (0.9%)
4	1 (0.1%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.0, 2.0, 43.0] 3.2 / 4.5
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 10.00] 0.65 / 1.46
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 14.00] 0.71 / 1.55
SELECT_n_of_aggr_func___SUM	
0	941 (94%)
1	41 (4.1%)
2	14 (1.4%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 10.00] 0.57 / 1.38
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 10.00] 0.66 / 1.51
SELECT_n_of_aggr_func___AVG	

Query parameter or value	Descriptive statistics
0	940 (94%)
1	48 (4.8%)
2	9 (0.9%)
3	3 (0.3%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 15.00] 2.73 / 2.35
FROM_n_of_join_paths	
1	342 (34%)
2	330 (33%)
3	328 (33%)
FROM_n_of_super_joins___FULL	
0	775 (78%)
1	202 (20%)
2	23 (2.3%)
FROM_n_of_super_joins___LEFT	
0	766 (77%)
1	208 (21%)
2	26 (2.6%)
FROM_n_of_super_joins___RIGHT	
0	779 (78%)
1	202 (20%)
2	19 (1.9%)
FROM_n_of_joins___INNER	
0	255 (26%)
1	275 (28%)
2	201 (20%)
3	135 (14%)
4	75 (7.5%)
5	35 (3.5%)
6	15 (1.5%)
7	6 (0.6%)
8	3 (0.3%)
FROM_n_of_joins___RIGHT	
0	236 (24%)
1	280 (28%)
2	211 (21%)
3	149 (15%)
4	66 (6.6%)
5	36 (3.6%)
6	16 (1.6%)
7	5 (0.5%)
9	1 (0.1%)
FROM_n_of_processed_rows	[5, 13,615,796, 45,887,231] 12,600,952 / 11,558,159
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.79 / 2.21
WHERE_n_of_attribs_of_type___character_varying	
0	468 (47%)
1	303 (30%)

Query parameter or value	Descriptive statistics
2	152 (15%)
3	59 (5.9%)
4	10 (1.0%)
5	7 (0.7%)
6	1 (0.1%)
WHERE_n_of_attribs_of_type___integer	
0	459 (46%)
1	310 (31%)
2	156 (16%)
3	56 (5.6%)
4	17 (1.7%)
5	1 (0.1%)
6	1 (0.1%)
WHERE_n_of_attribs_of_type___numeric	
0	652 (65%)
1	232 (23%)
2	96 (9.6%)
3	18 (1.8%)
4	1 (0.1%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type___character	
0	695 (70%)
1	236 (24%)
2	60 (6.0%)
3	8 (0.8%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type___date	
0	838 (84%)
1	137 (14%)
2	24 (2.4%)
3	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	625 (63%)
1	276 (28%)
2	85 (8.5%)
3	13 (1.3%)
4	1 (0.1%)
WHERE_n_of_connect_OR	[0.00, 1.00, 9.00] 1.36 / 1.67
WHERE_n_of_operators___between	
0	610 (61%)
1	284 (28%)
2	84 (8.4%)
3	18 (1.8%)
4	3 (0.3%)
5	1 (0.1%)
WHERE_n_of_operators___greater_or_less	

Query parameter or value	Descriptive statistics
0	419 (42%)
1	291 (29%)
2	174 (17%)
3	75 (7.5%)
4	30 (3.0%)
5	9 (0.9%)
6	2 (0.2%)
WHERE_n_of_operators___like	
0	848 (85%)
1	133 (13%)
2	17 (1.7%)
3	2 (0.2%)
WHERE_n_of_operators___in	
0	631 (63%)
1	251 (25%)
2	99 (9.9%)
3	18 (1.8%)
4	1 (0.1%)
WHERE_n_of_non_aggr_func___SQRT	
0	905 (91%)
1	88 (8.8%)
2	7 (0.7%)
WHERE_n_of_non_aggr_func___TRUNC	
0	970 (97%)
1	29 (2.9%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___ROUND	
0	958 (96%)
1	42 (4.2%)
WHERE_n_of_non_aggr_func___ABS	
0	957 (96%)
1	42 (4.2%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___FLOOR	
0	972 (97%)
1	28 (2.8%)
WHERE_n_of_non_aggr_func___YEAR	
0	972 (97%)
1	28 (2.8%)
WHERE_n_of_non_aggr_func___DOW	
0	970 (97%)
1	30 (3.0%)
WHERE_n_of_non_aggr_func___DAY	
0	982 (98%)
1	17 (1.7%)
2	1 (0.1%)

Query parameter or value	Descriptive statistics
WHERE_n_of_non_aggr_func___LOG	
0	948 (95%)
1	51 (5.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___MONTH	
0	973 (97%)
1	27 (2.7%)
WHERE_n_of_all_non_aggr_func	
0	696 (70%)
1	226 (23%)
2	60 (6.0%)
3	15 (1.5%)
4	2 (0.2%)
5	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.37 / 2.08
HAVING_n_of_main_predicates	
0	419 (42%)
1	195 (20%)
2	215 (22%)
3	171 (17%)
ORDER_BY_n_of_columns	
0	323 (32%)
1	377 (38%)
2	218 (22%)
3	54 (5.4%)
4	17 (1.7%)
5	7 (0.7%)
6	3 (0.3%)
8	1 (0.1%)
limit	[2, 492, 1,000] 495 / 290
offset	[0, 7, 1,000] 255 / 332