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

Scenario: 1-1jp__g__h0-3ss_h0-3nss

Scale factor: 0.01 GB

The 1000-query pack was generated on: 2023-11-05

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] 5.7 / 5.0
SELECT_n_of_non_aggr_func___RTRIM	
0	777 (78%)
1	196 (20%)
2	26 (2.6%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___DOW	
0	921 (92%)
1	78 (7.8%)
2	1 (0.1%)
SELECT_n_of_non_aggr_func___FLOOR	
0	969 (97%)
1	28 (2.8%)
2	3 (0.3%)
SELECT_n_of_non_aggr_func___ABS	

Query parameter or value	Descriptive statistics
0	966 (97%)
1	33 (3.3%)
2	1 (0.1%)
SELECT_n_of_non_aggr_func___LOWER	
0	787 (79%)
1	191 (19%)
2	21 (2.1%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	779 (78%)
1	201 (20%)
2	18 (1.8%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___UPPER	
0	782 (78%)
1	189 (19%)
2	23 (2.3%)
3	6 (0.6%)
SELECT_n_of_non_aggr_func___SQRT	
0	889 (89%)
1	105 (11%)
2	6 (0.6%)
SELECT_n_of_non_aggr_func___MONTH	
0	889 (89%)
1	94 (9.4%)
2	15 (1.5%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___YEAR	
0	912 (91%)
1	73 (7.3%)
2	12 (1.2%)
3	2 (0.2%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___LTRIM	
0	788 (79%)
1	184 (18%)
2	26 (2.6%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___TRUNC	
0	955 (96%)
1	45 (4.5%)
SELECT_n_of_non_aggr_func___DAY	
0	946 (95%)
1	51 (5.1%)
2	3 (0.3%)
SELECT_n_of_non_aggr_func___LOG	

Query parameter or value	Descriptive statistics
0	939 (94%)
1	56 (5.6%)
2	4 (0.4%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___ROUND	
0	969 (97%)
1	31 (3.1%)
SELECT_n_of_all_non_aggr_func	[0.00, 1.00, 13.00] 1.93 / 1.98
SELECT_n_of_aggr_func___COUNT_DISTINCT	
0	773 (77%)
1	110 (11%)
2	59 (5.9%)
3	26 (2.6%)
4	21 (2.1%)
5	10 (1.0%)
7	1 (0.1%)
SELECT_n_of_aggr_func___COUNT	
0	773 (77%)
1	99 (9.9%)
2	70 (7.0%)
3	33 (3.3%)
4	14 (1.4%)
5	10 (1.0%)
6	1 (0.1%)
SELECT_n_of_aggr_func___MIN	
0	765 (77%)
1	114 (11%)
2	69 (6.9%)
3	38 (3.8%)
4	4 (0.4%)
5	7 (0.7%)
6	3 (0.3%)
SELECT_n_of_aggr_func___SUM	
0	949 (95%)
1	37 (3.7%)
2	10 (1.0%)
3	4 (0.4%)
SELECT_n_of_aggr_func___MAX	
0	783 (78%)
1	109 (11%)
2	55 (5.5%)
3	25 (2.5%)
4	16 (1.6%)
5	11 (1.1%)
6	1 (0.1%)
SELECT_n_of_aggr_func___AVG	

Query parameter or value	Descriptive statistics
0	951 (95%)
1	40 (4.0%)
2	7 (0.7%)
3	2 (0.2%)
SELECT_n_of_all_aggr_func	
0	189 (19%)
1	264 (26%)
2	264 (26%)
3	140 (14%)
4	68 (6.8%)
5	56 (5.6%)
6	17 (1.7%)
7	2 (0.2%)
FROM_n_of_join_paths	
1	1,000 (100%)
FROM_n_of_joins___RIGHT	
0	485 (49%)
1	326 (33%)
2	137 (14%)
3	44 (4.4%)
4	8 (0.8%)
FROM_n_of_joins___INNER	
0	495 (50%)
1	312 (31%)
2	151 (15%)
3	35 (3.5%)
4	7 (0.7%)
FROM_n_of_processed_rows	[5, 15,000, 76,705] 29,574 / 31,385
WHERE_n_of_predicates	
0	240 (24%)
1	238 (24%)
2	208 (21%)
3	150 (15%)
4	100 (10%)
5	56 (5.6%)
6	8 (0.8%)
WHERE_n_of_attribs_of_type___character_varying	
0	536 (54%)
1	323 (32%)
2	118 (12%)
3	21 (2.1%)
4	1 (0.1%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type___integer	
0	588 (59%)
1	282 (28%)

Query parameter or value	Descriptive statistics
2	113 (11%)
3	16 (1.6%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type___numeric	
0	732 (73%)
1	221 (22%)
2	41 (4.1%)
3	6 (0.6%)
WHERE_n_of_attribs_of_type___character	
0	820 (82%)
1	163 (16%)
2	15 (1.5%)
3	2 (0.2%)
WHERE_n_of_attribs_of_type___date	
0	891 (89%)
1	99 (9.9%)
2	9 (0.9%)
3	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	733 (73%)
1	224 (22%)
2	38 (3.8%)
3	5 (0.5%)
WHERE_n_of_connect_OR	
0	610 (61%)
1	203 (20%)
2	110 (11%)
3	62 (6.2%)
4	15 (1.5%)
WHERE_n_of_operators___greater_or_less	
0	527 (53%)
1	303 (30%)
2	122 (12%)
3	41 (4.1%)
4	7 (0.7%)
WHERE_n_of_operators___in	
0	707 (71%)
1	244 (24%)
2	43 (4.3%)
3	6 (0.6%)
WHERE_n_of_operators___between	
0	724 (72%)
1	234 (23%)
2	36 (3.6%)
3	5 (0.5%)
4	1 (0.1%)

Query parameter or value	Descriptive statistics
WHERE_n_of_operators___like	
0	873 (87%)
1	118 (12%)
2	9 (0.9%)
WHERE_n_of_non_aggr_func___FLOOR	
0	973 (97%)
1	27 (2.7%)
WHERE_n_of_non_aggr_func___SQRT	
0	933 (93%)
1	63 (6.3%)
2	4 (0.4%)
WHERE_n_of_non_aggr_func___ABS	
0	980 (98%)
1	20 (2.0%)
WHERE_n_of_non_aggr_func___MONTH	
0	980 (98%)
1	20 (2.0%)
WHERE_n_of_non_aggr_func___YEAR	
0	982 (98%)
1	18 (1.8%)
WHERE_n_of_non_aggr_func___TRUNC	
0	980 (98%)
1	20 (2.0%)
WHERE_n_of_non_aggr_func___LOG	
0	958 (96%)
1	41 (4.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___ROUND	
0	976 (98%)
1	24 (2.4%)
WHERE_n_of_non_aggr_func___DAY	
0	990 (99%)
1	10 (1.0%)
WHERE_n_of_non_aggr_func___DOW	
0	986 (99%)
1	14 (1.4%)
WHERE_n_of_all_non_aggr_func	
0	781 (78%)
1	177 (18%)
2	36 (3.6%)
3	6 (0.6%)
GROUP_BY_n_of_columns	
0	251 (25%)
1	280 (28%)
2	207 (21%)
3	151 (15%)

Query parameter or value	Descriptive statistics
4	77 (7.7%)
5	31 (3.1%)
6	3 (0.3%)
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.40 / 2.61
HAVING_n_of_main_predicates___non_scalar_subquery	
0	442 (44%)
1	196 (20%)
2	190 (19%)
3	172 (17%)
HAVING_n_of_main_predicates___scalar_subquery	
0	429 (43%)
1	192 (19%)
2	182 (18%)
3	197 (20%)
HAVING_n_of_subqueries___non_scalar_subquery	
0	442 (44%)
1	196 (20%)
2	190 (19%)
3	172 (17%)
HAVING_n_of_subqueries___scalar_subquery	[0.0, 1.0, 20.0] 2.3 / 3.3
HAVING_n_of_processed_rows_by_subqueries	[0, 8,130, 1,366,000] 104,965 / 201,788
ORDER_BY_n_of_columns	
0	287 (29%)
1	484 (48%)
2	194 (19%)
3	28 (2.8%)
4	6 (0.6%)
5	1 (0.1%)
limit	[2, 496, 1,000] 489 / 289
offset	[0, 58, 1,000] 257 / 318