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

Scenario: 1-3jp_rjjp__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, 6, 102] 10 / 12
SELECT_n_of_non_aggr_func___LOWER	
0	704 (70%)
1	234 (23%)
2	46 (4.6%)
3	13 (1.3%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	681 (68%)
1	243 (24%)
2	57 (5.7%)
3	18 (1.8%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___LOG	

Query parameter or value	Descriptive statistics
0	916 (92%)
1	74 (7.4%)
2	7 (0.7%)
3	2 (0.2%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___LTRIM	
0	689 (69%)
1	240 (24%)
2	53 (5.3%)
3	11 (1.1%)
4	4 (0.4%)
5	3 (0.3%)
SELECT_n_of_non_aggr_func___MONTH	
0	845 (85%)
1	119 (12%)
2	22 (2.2%)
3	10 (1.0%)
4	1 (0.1%)
5	3 (0.3%)
SELECT_n_of_non_aggr_func___ROUND	
0	948 (95%)
1	45 (4.5%)
2	7 (0.7%)
SELECT_n_of_non_aggr_func___SQRT	
0	840 (84%)
1	127 (13%)
2	21 (2.1%)
3	10 (1.0%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___UPPER	
0	708 (71%)
1	224 (22%)
2	52 (5.2%)
3	10 (1.0%)
4	4 (0.4%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___YEAR	
0	864 (86%)
1	103 (10%)
2	23 (2.3%)
3	6 (0.6%)
4	3 (0.3%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___FLOOR	
0	948 (95%)

Query parameter or value	Descriptive statistics
1	47 (4.7%)
2	4 (0.4%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___RTRIM	
0	726 (73%)
1	217 (22%)
2	41 (4.1%)
3	11 (1.1%)
4	2 (0.2%)
5	2 (0.2%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___DAY	
0	908 (91%)
1	79 (7.9%)
2	13 (1.3%)
SELECT_n_of_non_aggr_func___DOW	
0	884 (88%)
1	89 (8.9%)
2	21 (2.1%)
3	4 (0.4%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func___TRUNC	
0	949 (95%)
1	48 (4.8%)
2	3 (0.3%)
SELECT_n_of_non_aggr_func___ABS	
0	951 (95%)
1	43 (4.3%)
2	5 (0.5%)
3	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.00, 2.00, 35.00] 3.13 / 4.00
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 10.00] 0.67 / 1.49
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 10.00] 0.59 / 1.42
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 10.00] 0.65 / 1.49
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 10.00] 0.64 / 1.46
SELECT_n_of_aggr_func___SUM	
0	936 (94%)
1	44 (4.4%)
2	15 (1.5%)
3	5 (0.5%)
SELECT_n_of_aggr_func___AVG	
0	945 (95%)
1	40 (4.0%)
2	12 (1.2%)
3	2 (0.2%)
4	1 (0.1%)

Query parameter or value	Descriptive statistics
SELECT_n_of_all_aggr_func	[0.00, 2.00, 11.00] 2.71 / 2.33
FROM_n_of_join_paths	
1	336 (34%)
2	321 (32%)
3	343 (34%)
FROM_n_of_super_joins___FULL	
0	777 (78%)
1	203 (20%)
2	20 (2.0%)
FROM_n_of_super_joins___LEFT	
0	758 (76%)
1	223 (22%)
2	19 (1.9%)
FROM_n_of_super_joins___RIGHT	
0	772 (77%)
1	214 (21%)
2	14 (1.4%)
FROM_n_of_joins___INNER	
0	235 (24%)
1	240 (24%)
2	217 (22%)
3	176 (18%)
4	83 (8.3%)
5	34 (3.4%)
6	12 (1.2%)
7	2 (0.2%)
8	1 (0.1%)
FROM_n_of_joins___RIGHT	
0	269 (27%)
1	247 (25%)
2	214 (21%)
3	132 (13%)
4	80 (8.0%)
5	35 (3.5%)
6	15 (1.5%)
7	8 (0.8%)
FROM_n_of_processed_rows	[5, 68,333, 228,580] 66,588 / 60,237
WHERE_n_of_predicates	[0.00, 3.00, 10.00] 2.82 / 2.21
WHERE_n_of_attribs_of_type___character_varying	
0	469 (47%)
1	285 (29%)
2	154 (15%)
3	67 (6.7%)
4	21 (2.1%)
5	4 (0.4%)
WHERE_n_of_attribs_of_type___integer	

Query parameter or value	Descriptive statistics
0	433 (43%)
1	352 (35%)
2	153 (15%)
3	34 (3.4%)
4	19 (1.9%)
5	9 (0.9%)
WHERE_n_of_attribs_of_type___character	
0	722 (72%)
1	214 (21%)
2	57 (5.7%)
3	7 (0.7%)
WHERE_n_of_attribs_of_type___date	
0	826 (83%)
1	150 (15%)
2	21 (2.1%)
3	3 (0.3%)
WHERE_n_of_attribs_of_type___numeric	
0	634 (63%)
1	267 (27%)
2	78 (7.8%)
3	19 (1.9%)
4	2 (0.2%)
WHERE_n_of_pkey_attribs	
0	615 (62%)
1	301 (30%)
2	65 (6.5%)
3	18 (1.8%)
4	1 (0.1%)
WHERE_n_of_connect_OR	
0	444 (44%)
1	181 (18%)
2	137 (14%)
3	111 (11%)
4	61 (6.1%)
5	31 (3.1%)
6	25 (2.5%)
7	8 (0.8%)
8	2 (0.2%)
WHERE_n_of_operators___between	
0	619 (62%)
1	261 (26%)
2	91 (9.1%)
3	24 (2.4%)
4	5 (0.5%)
WHERE_n_of_operators___greater_or_less	
0	416 (42%)

Query parameter or value	Descriptive statistics
1	298 (30%)
2	189 (19%)
3	54 (5.4%)
4	32 (3.2%)
5	10 (1.0%)
6	1 (0.1%)
WHERE_n_of_operators___in	
0	603 (60%)
1	281 (28%)
2	93 (9.3%)
3	22 (2.2%)
4	1 (0.1%)
WHERE_n_of_operators___like	
0	830 (83%)
1	153 (15%)
2	17 (1.7%)
WHERE_n_of_non_aggr_func___DAY	
0	970 (97%)
1	28 (2.8%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___TRUNC	
0	968 (97%)
1	30 (3.0%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___ROUND	
0	976 (98%)
1	24 (2.4%)
WHERE_n_of_non_aggr_func___SQRT	
0	903 (90%)
1	88 (8.8%)
2	9 (0.9%)
WHERE_n_of_non_aggr_func___YEAR	
0	974 (97%)
1	26 (2.6%)
WHERE_n_of_non_aggr_func___LOG	
0	930 (93%)
1	67 (6.7%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func___ABS	
0	961 (96%)
1	39 (3.9%)
WHERE_n_of_non_aggr_func___MONTH	
0	965 (97%)
1	34 (3.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___FLOOR	

Query parameter or value	Descriptive statistics
0	971 (97%)
1	28 (2.8%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___DOW	
0	975 (98%)
1	24 (2.4%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	691 (69%)
1	217 (22%)
2	72 (7.2%)
3	16 (1.6%)
4	3 (0.3%)
5	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.27 / 2.11
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.34 / 2.67
HAVING_n_of_main_predicates___non_scalar_subquery	
0	460 (46%)
1	184 (18%)
2	173 (17%)
3	183 (18%)
HAVING_n_of_main_predicates___scalar_subquery	
0	453 (45%)
1	182 (18%)
2	166 (17%)
3	199 (20%)
HAVING_n_of_subqueries___non_scalar_subquery	
0	460 (46%)
1	184 (18%)
2	173 (17%)
3	183 (18%)
HAVING_n_of_subqueries___scalar_subquery	[0.0, 1.0, 20.0] 2.4 / 3.6
HAVING_n_of_processed_rows_by_subqueries	[0, 33,330, 3,174,825] 239,753 / 434,898
ORDER_BY_n_of_columns	
0	304 (30%)
1	381 (38%)
2	235 (24%)
3	60 (6.0%)
4	12 (1.2%)
5	6 (0.6%)
6	1 (0.1%)
7	1 (0.1%)
limit	[1, 473, 997] 480 / 292
offset	[0, 0, 999] 246 / 322