

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

Scenario: 1-6jp_rjpp__g__h0-3ss_h0-3nss

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, 8, 172] 15 / 22
SELECT_n_of_non_aggr_func__LOWER	[0.00, 0.00, 11.00] 0.61 / 1.08
SELECT_n_of_non_aggr_func__DOW	
0	840 (84%)
1	113 (11%)
2	26 (2.6%)
3	6 (0.6%)
4	8 (0.8%)
5	4 (0.4%)
6	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__MONTH	
0	791 (79%)
1	140 (14%)
2	34 (3.4%)

Query parameter or value	Descriptive statistics
3	10 (1.0%)
4	8 (0.8%)
5	7 (0.7%)
6	4 (0.4%)
7	3 (0.3%)
8	3 (0.3%)
SELECT_n_of_non_aggr_func__RTRIM	[0.00, 0.00, 10.00] 0.65 / 1.08
SELECT_n_of_non_aggr_func__SUBSTR	
0	622 (62%)
1	238 (24%)
10	1 (0.1%)
2	84 (8.4%)
3	24 (2.4%)
4	12 (1.2%)
5	11 (1.1%)
6	6 (0.6%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func__YEAR	
0	798 (80%)
1	149 (15%)
2	23 (2.3%)
3	12 (1.2%)
4	10 (1.0%)
5	4 (0.4%)
6	2 (0.2%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func__DAY	
0	870 (87%)
1	94 (9.4%)
2	20 (2.0%)
3	10 (1.0%)
4	5 (0.5%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__LTRIM	[0.00, 0.00, 10.00] 0.65 / 1.15
SELECT_n_of_non_aggr_func__UPPER	[0.00, 0.00, 10.00] 0.68 / 1.18
SELECT_n_of_non_aggr_func__ABS	
0	907 (91%)
1	77 (7.7%)
2	11 (1.1%)
3	4 (0.4%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__FLOOR	
0	931 (93%)
1	51 (5.1%)
2	15 (1.5%)
3	3 (0.3%)

Query parameter or value	Descriptive statistics
SELECT__n_of__non__aggr__func__LOG	
0	843 (84%)
1	117 (12%)
2	24 (2.4%)
3	13 (1.3%)
4	2 (0.2%)
5	1 (0.1%)
SELECT__n_of__non__aggr__func__ROUND	
0	908 (91%)
1	71 (7.1%)
2	14 (1.4%)
3	5 (0.5%)
4	2 (0.2%)
SELECT__n_of__non__aggr__func__SQRT	
0	809 (81%)
1	134 (13%)
2	25 (2.5%)
3	15 (1.5%)
4	11 (1.1%)
5	4 (0.4%)
6	1 (0.1%)
8	1 (0.1%)
SELECT__n_of__non__aggr__func__TRUNC	
0	924 (92%)
1	60 (6.0%)
2	10 (1.0%)
3	4 (0.4%)
4	1 (0.1%)
5	1 (0.1%)
SELECT__n_of__all__non__aggr__func	[0, 3, 69] 5 / 9
SELECT__n_of__aggr__func__MIN	[0.00, 0.00, 13.00] 0.88 / 1.94
SELECT__n_of__aggr__func__MAX	[0.00, 0.00, 11.00] 0.84 / 1.91
SELECT__n_of__aggr__func__COUNT	[0.00, 0.00, 13.00] 0.86 / 2.03
SELECT__n_of__aggr__func__COUNT_DISTINCT	[0.00, 0.00, 12.00] 0.78 / 1.86
SELECT__n_of__aggr__func__SUM	
0	935 (94%)
1	45 (4.5%)
2	13 (1.3%)
3	6 (0.6%)
5	1 (0.1%)
SELECT__n_of__aggr__func__AVG	
0	935 (94%)
1	41 (4.1%)
2	15 (1.5%)
3	5 (0.5%)
4	4 (0.4%)

Query parameter or value	Descriptive statistics
SELECT_n_of_all_aggr_func	[0.0, 3.0, 14.0] 3.6 / 3.1
FROM_n_of_join_paths	
1	171 (17%)
2	182 (18%)
3	175 (18%)
4	157 (16%)
5	159 (16%)
6	156 (16%)
FROM_n_of_super_joins___FULL	
0	569 (57%)
1	294 (29%)
2	117 (12%)
3	18 (1.8%)
4	2 (0.2%)
FROM_n_of_super_joins___LEFT	
0	533 (53%)
1	322 (32%)
2	115 (12%)
3	27 (2.7%)
4	3 (0.3%)
FROM_n_of_super_joins___RIGHT	
0	560 (56%)
1	296 (30%)
2	108 (11%)
3	31 (3.1%)
4	5 (0.5%)
FROM_n_of_joins___INNER	[0.00, 3.00, 11.00] 3.16 / 2.42
FROM_n_of_joins___RIGHT	[0.00, 3.00, 13.00] 3.26 / 2.57
FROM_n_of_processed_rows	[5, 44,025,764, 221,100,022] 57,845,081 / 49,287,417
WHERE_n_of_predicates	[0.00, 3.00, 13.00] 3.73 / 2.96
WHERE_n_of_attrbs_of_type___character_varying	
0	382 (38%)
1	275 (28%)
2	196 (20%)
3	98 (9.8%)
4	24 (2.4%)
5	19 (1.9%)
6	5 (0.5%)
7	1 (0.1%)
WHERE_n_of_attrbs_of_type___integer	
0	386 (39%)
1	276 (28%)
2	181 (18%)
3	92 (9.2%)
4	46 (4.6%)
5	17 (1.7%)

Query parameter or value	Descriptive statistics
6	1 (0.1%)
7	1 (0.1%)
WHERE_n_of_attribs_of_type_numeric	
0	582 (58%)
1	263 (26%)
2	106 (11%)
3	38 (3.8%)
4	7 (0.7%)
5	2 (0.2%)
6	2 (0.2%)
WHERE_n_of_attribs_of_type_character	
0	661 (66%)
1	247 (25%)
2	73 (7.3%)
3	18 (1.8%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type_date	
0	795 (80%)
1	163 (16%)
2	33 (3.3%)
3	9 (0.9%)
WHERE_n_of_pkey_attribs	
0	533 (53%)
1	300 (30%)
2	122 (12%)
3	34 (3.4%)
4	7 (0.7%)
5	4 (0.4%)
WHERE_n_of_connect_OR	[0.00, 1.00, 11.00] 2.13 / 2.44
WHERE_n_of_operators_between	
0	549 (55%)
1	271 (27%)
2	127 (13%)
3	38 (3.8%)
4	12 (1.2%)
5	3 (0.3%)
WHERE_n_of_operators_greater_or_less	
0	339 (34%)
1	283 (28%)
2	184 (18%)
3	111 (11%)
4	49 (4.9%)
5	22 (2.2%)
6	7 (0.7%)
7	5 (0.5%)
WHERE_n_of_operators_like	

Query parameter or value	Descriptive statistics
0	786 (79%)
1	187 (19%)
2	27 (2.7%)
WHERE_n_of_operators_in	
0	543 (54%)
1	280 (28%)
2	111 (11%)
3	46 (4.6%)
4	17 (1.7%)
5	3 (0.3%)
WHERE_n_of_non_aggr_func_FLOOR	
0	958 (96%)
1	41 (4.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_SQRT	
0	891 (89%)
1	104 (10%)
2	5 (0.5%)
WHERE_n_of_non_aggr_func_DOW	
0	962 (96%)
1	37 (3.7%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_LOG	
0	906 (91%)
1	89 (8.9%)
2	3 (0.3%)
3	2 (0.2%)
WHERE_n_of_non_aggr_func_TRUNC	
0	956 (96%)
1	42 (4.2%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func_YEAR	
0	968 (97%)
1	31 (3.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_MONTH	
0	963 (96%)
1	36 (3.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_ROUND	
0	961 (96%)
1	37 (3.7%)
2	1 (0.1%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func_DAY	
0	962 (96%)

Query parameter or value	Descriptive statistics
1	38 (3.8%)
WHERE_n_of_non_aggr_func__ABS	
0	952 (95%)
1	46 (4.6%)
2	2 (0.2%)
WHERE_n_of_all_non_aggr_func	
0	625 (63%)
1	246 (25%)
2	97 (9.7%)
3	25 (2.5%)
4	6 (0.6%)
5	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 13.00] 3.02 / 2.85
HAVING_n_of_main_predicates	[0.00, 3.00, 9.00] 3.23 / 2.57
HAVING_n_of_main_predicates__non_scalar_subquery	
0	455 (46%)
1	200 (20%)
2	176 (18%)
3	169 (17%)
HAVING_n_of_main_predicates__scalar_subquery	
0	452 (45%)
1	174 (17%)
2	202 (20%)
3	172 (17%)
HAVING_n_of_subqueries__non_scalar_subquery	
0	455 (46%)
1	200 (20%)
2	176 (18%)
3	169 (17%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 18.0] 2.2 / 3.2
HAVING_n_of_processed_rows_by_subqueries	[0, 51,752,555, 3,229,205,526] 193,121,108 / 333,295,065
ORDER_BY_n_of_columns	
0	305 (31%)
1	341 (34%)
2	226 (23%)
3	96 (9.6%)
4	22 (2.2%)
5	4 (0.4%)
6	5 (0.5%)
8	1 (0.1%)
limit	[1, 529, 1,000] 513 / 281
offset	[0, 0, 999] 242 / 322