

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

Scenario: 1-6jp_ujpp__g__h0-3ss_h0-3nss

Scale factor: 10 GB

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

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, 198] 15 / 24
SELECT_n_of_non_aggr_func__RTRIM	[0.00, 0.00, 9.00] 0.58 / 0.99
SELECT_n_of_non_aggr_func__UPPER	
0	619 (62%)
1	269 (27%)
2	76 (7.6%)
3	18 (1.8%)
4	10 (1.0%)
5	2 (0.2%)
6	4 (0.4%)
7	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__LTRIM	
0	622 (62%)
1	266 (27%)

Query parameter or value	Descriptive statistics
2	72 (7.2%)
3	25 (2.5%)
4	7 (0.7%)
5	5 (0.5%)
6	1 (0.1%)
8	2 (0.2%)
SELECT__n_of__non__aggr__func__SUBSTR	
0	615 (62%)
1	261 (26%)
2	87 (8.7%)
3	23 (2.3%)
4	7 (0.7%)
5	2 (0.2%)
6	4 (0.4%)
9	1 (0.1%)
SELECT__n_of__non__aggr__func__DOW	
0	856 (86%)
1	113 (11%)
2	17 (1.7%)
3	8 (0.8%)
4	1 (0.1%)
5	5 (0.5%)
SELECT__n_of__non__aggr__func__SQRT	
0	825 (83%)
1	127 (13%)
2	29 (2.9%)
3	12 (1.2%)
4	4 (0.4%)
5	2 (0.2%)
6	1 (0.1%)
SELECT__n_of__non__aggr__func__LOWER	
0	630 (63%)
1	240 (24%)
2	88 (8.8%)
3	24 (2.4%)
4	12 (1.2%)
5	3 (0.3%)
6	1 (0.1%)
7	2 (0.2%)
SELECT__n_of__non__aggr__func__DAY	
0	903 (90%)
1	71 (7.1%)
2	20 (2.0%)
3	4 (0.4%)
4	2 (0.2%)
SELECT__n_of__non__aggr__func__MONTH	

Query parameter or value	Descriptive statistics
0	819 (82%)
1	125 (13%)
2	30 (3.0%)
3	17 (1.7%)
4	6 (0.6%)
6	1 (0.1%)
7	1 (0.1%)
9	1 (0.1%)
SELECT__n_of__non__aggr__func__FLOOR	
0	931 (93%)
1	55 (5.5%)
2	9 (0.9%)
3	4 (0.4%)
4	1 (0.1%)
SELECT__n_of__non__aggr__func__LOG	
0	867 (87%)
1	105 (11%)
2	18 (1.8%)
3	4 (0.4%)
4	3 (0.3%)
5	2 (0.2%)
6	1 (0.1%)
SELECT__n_of__non__aggr__func__ROUND	
0	925 (93%)
1	66 (6.6%)
2	4 (0.4%)
3	4 (0.4%)
5	1 (0.1%)
SELECT__n_of__non__aggr__func__TRUNC	
0	918 (92%)
1	71 (7.1%)
2	9 (0.9%)
3	2 (0.2%)
SELECT__n_of__non__aggr__func__YEAR	
0	848 (85%)
1	114 (11%)
2	19 (1.9%)
3	12 (1.2%)
4	4 (0.4%)
5	1 (0.1%)
6	2 (0.2%)
SELECT__n_of__non__aggr__func__ABS	
0	929 (93%)
1	64 (6.4%)
2	5 (0.5%)
3	2 (0.2%)

Query parameter or value	Descriptive statistics
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 74.0] 4.5 / 6.3
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 11.00] 0.74 / 1.73
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 13.00] 0.81 / 1.79
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 12.00] 1.01 / 2.12
SELECT_n_of_aggr_func__SUM	
0	938 (94%)
1	40 (4.0%)
2	16 (1.6%)
3	4 (0.4%)
4	2 (0.2%)
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 12.00] 0.88 / 1.86
SELECT_n_of_aggr_func__AVG	
0	929 (93%)
1	38 (3.8%)
2	17 (1.7%)
3	12 (1.2%)
4	4 (0.4%)
SELECT_n_of_all_aggr_func	[0.00, 3.00, 13.00] 3.66 / 3.00
FROM_n_of_join_paths	
1	158 (16%)
2	158 (16%)
3	182 (18%)
4	157 (16%)
5	173 (17%)
6	172 (17%)
FROM_n_of_super_joins__LEFT	
0	545 (55%)
1	321 (32%)
2	104 (10%)
3	30 (3.0%)
FROM_n_of_super_joins__FULL	
0	522 (52%)
1	316 (32%)
2	134 (13%)
3	26 (2.6%)
4	2 (0.2%)
FROM_n_of_super_joins__RIGHT	
0	558 (56%)
1	293 (29%)
2	118 (12%)
3	30 (3.0%)
4	1 (0.1%)
FROM_n_of_joins__INNER	[0.00, 3.00, 13.00] 3.37 / 2.54
FROM_n_of_joins__RIGHT	[0.00, 3.00, 12.00] 3.41 / 2.42
FROM_n_of_processed_rows	[5, 94,482,757, 449,382,022] 120,526,821 / 97,239,315
WHERE_n_of_predicates	[0.0, 3.0, 13.0] 3.9 / 3.0

Query parameter or value	Descriptive statistics
WHERE_n_of_attribs_of_type__character_varying	
0	362 (36%)
1	293 (29%)
2	194 (19%)
3	88 (8.8%)
4	47 (4.7%)
5	12 (1.2%)
6	4 (0.4%)
WHERE_n_of_attribs_of_type__integer	
0	394 (39%)
1	267 (27%)
2	185 (19%)
3	96 (9.6%)
4	35 (3.5%)
5	17 (1.7%)
6	5 (0.5%)
7	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	750 (75%)
1	203 (20%)
2	42 (4.2%)
3	5 (0.5%)
WHERE_n_of_attribs_of_type__character	
0	637 (64%)
1	262 (26%)
2	77 (7.7%)
3	20 (2.0%)
4	4 (0.4%)
WHERE_n_of_attribs_of_type__numeric	
0	562 (56%)
1	267 (27%)
2	126 (13%)
3	32 (3.2%)
4	9 (0.9%)
5	3 (0.3%)
6	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	547 (55%)
1	303 (30%)
2	114 (11%)
3	29 (2.9%)
4	4 (0.4%)
5	3 (0.3%)
WHERE_n_of_connect_OR	[0.00, 1.00, 12.00] 2.25 / 2.48
WHERE_n_of_operators__greater_or_less	
0	341 (34%)

Query parameter or value	Descriptive statistics
1	263 (26%)
2	170 (17%)
3	123 (12%)
4	58 (5.8%)
5	30 (3.0%)
6	10 (1.0%)
7	4 (0.4%)
8	1 (0.1%)
WHERE_n_of_operators__in	
0	524 (52%)
1	278 (28%)
2	148 (15%)
3	40 (4.0%)
4	10 (1.0%)
WHERE_n_of_operators__between	
0	542 (54%)
1	298 (30%)
2	105 (11%)
3	38 (3.8%)
4	11 (1.1%)
5	5 (0.5%)
6	1 (0.1%)
WHERE_n_of_operators__like	
0	774 (77%)
1	197 (20%)
2	27 (2.7%)
3	2 (0.2%)
WHERE_n_of_non_aggr_func__LOG	
0	921 (92%)
1	76 (7.6%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func__FLOOR	
0	957 (96%)
1	42 (4.2%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__TRUNC	
0	954 (95%)
1	46 (4.6%)
WHERE_n_of_non_aggr_func__SQRT	
0	884 (88%)
1	107 (11%)
2	9 (0.9%)
WHERE_n_of_non_aggr_func__YEAR	
0	956 (96%)
1	44 (4.4%)
WHERE_n_of_non_aggr_func__ROUND	

Query parameter or value	Descriptive statistics
0	961 (96%)
1	38 (3.8%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DOW	
0	953 (95%)
1	41 (4.1%)
2	6 (0.6%)
WHERE_n_of_non_aggr_func__ABS	
0	973 (97%)
1	26 (2.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DAY	
0	962 (96%)
1	38 (3.8%)
WHERE_n_of_non_aggr_func__MONTH	
0	966 (97%)
1	34 (3.4%)
WHERE_n_of_all_non_aggr_func	
0	627 (63%)
1	253 (25%)
2	86 (8.6%)
3	28 (2.8%)
4	5 (0.5%)
5	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 3.00, 13.00] 3.33 / 2.90
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.44 / 2.52
HAVING_n_of_main_predicates__non_scalar_subquery	
0	424 (42%)
1	195 (20%)
2	197 (20%)
3	184 (18%)
HAVING_n_of_main_predicates__scalar_subquery	
0	441 (44%)
1	162 (16%)
2	195 (20%)
3	202 (20%)
HAVING_n_of_subqueries__non_scalar_subquery	
0	424 (42%)
1	195 (20%)
2	197 (20%)
3	184 (18%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 27.0] 2.5 / 3.8
HAVING_n_of_processed_rows_by_subqueries	[0, 155,360,205, 5,499,389,352] 424,438,726 / 693,645,939
ORDER_BY_n_of_columns	
0	298 (30%)

Query parameter or value	Descriptive statistics
<i>1</i>	333 (33%)
<i>2</i>	250 (25%)
<i>3</i>	96 (9.6%)
<i>4</i>	13 (1.3%)
<i>5</i>	5 (0.5%)
<i>6</i>	3 (0.3%)
<i>7</i>	1 (0.1%)
<i>8</i>	1 (0.1%)
limit	[1, 511, 1,000] 502 / 296
offset	[0, 0, 999] 230 / 315