

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

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

Scale factor: 2000 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, 117] 11 / 16
SELECT_n_of_non_aggr_func__LOWER	
0	712 (71%)
1	214 (21%)
2	53 (5.3%)
3	18 (1.8%)
4	1 (0.1%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func__SQRT	
0	861 (86%)
1	109 (11%)
2	19 (1.9%)
3	8 (0.8%)
4	2 (0.2%)
5	1 (0.1%)

Query parameter or value	Descriptive statistics
SELECT__n_of__non_aggr_func__SUBSTR	
0	669 (67%)
1	253 (25%)
2	50 (5.0%)
3	12 (1.2%)
4	8 (0.8%)
5	4 (0.4%)
6	3 (0.3%)
7	1 (0.1%)
SELECT__n_of__non_aggr_func__LOG	
0	901 (90%)
1	86 (8.6%)
2	7 (0.7%)
3	3 (0.3%)
4	1 (0.1%)
5	1 (0.1%)
7	1 (0.1%)
SELECT__n_of__non_aggr_func__RTRIM	
0	713 (71%)
1	213 (21%)
2	51 (5.1%)
3	16 (1.6%)
4	5 (0.5%)
5	1 (0.1%)
6	1 (0.1%)
SELECT__n_of__non_aggr_func__UPPER	
0	702 (70%)
1	226 (23%)
2	55 (5.5%)
3	12 (1.2%)
4	2 (0.2%)
5	2 (0.2%)
6	1 (0.1%)
SELECT__n_of__non_aggr_func__LTRIM	
0	691 (69%)
1	232 (23%)
2	56 (5.6%)
3	18 (1.8%)
4	2 (0.2%)
6	1 (0.1%)
SELECT__n_of__non_aggr_func__DAY	
0	907 (91%)
1	85 (8.5%)
2	5 (0.5%)
3	3 (0.3%)
SELECT__n_of__non_aggr_func__MONTH	

Query parameter or value	Descriptive statistics
0	847 (85%)
1	115 (12%)
2	23 (2.3%)
3	7 (0.7%)
4	3 (0.3%)
5	3 (0.3%)
6	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__ROUND	
0	941 (94%)
1	54 (5.4%)
2	3 (0.3%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__TRUNC	
0	939 (94%)
1	55 (5.5%)
2	4 (0.4%)
3	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	1 (0.1%)
5	3 (0.3%)
SELECT_n_of_non_aggr_func__ABS	
0	936 (94%)
1	55 (5.5%)
2	8 (0.8%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__DOW	
0	878 (88%)
1	96 (9.6%)
2	16 (1.6%)
3	8 (0.8%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__FLOOR	
0	940 (94%)
1	56 (5.6%)
2	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.0, 2.0, 40.0] 3.3 / 4.7
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 9.00] 0.57 / 1.31
SELECT_n_of_aggr_func__MAX	

Query parameter or value	Descriptive statistics
0	762 (76%)
1	90 (9.0%)
2	56 (5.6%)
3	33 (3.3%)
4	25 (2.5%)
5	18 (1.8%)
6	3 (0.3%)
7	8 (0.8%)
8	5 (0.5%)
SELECT_n_of_aggr_func__MIN	
0	743 (74%)
1	84 (8.4%)
2	60 (6.0%)
3	36 (3.6%)
4	32 (3.2%)
5	16 (1.6%)
6	11 (1.1%)
7	14 (1.4%)
8	4 (0.4%)
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 11.00] 0.67 / 1.53
SELECT_n_of_aggr_func__SUM	
0	943 (94%)
1	42 (4.2%)
2	11 (1.1%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_aggr_func__AVG	
0	945 (95%)
1	35 (3.5%)
2	13 (1.3%)
3	6 (0.6%)
4	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 12.00] 2.72 / 2.31
FROM_n_of_join_paths	
1	326 (33%)
2	334 (33%)
3	340 (34%)
FROM_n_of_super_joins__LEFT	
0	761 (76%)
1	217 (22%)
2	22 (2.2%)
FROM_n_of_super_joins__RIGHT	
0	786 (79%)
1	192 (19%)
2	22 (2.2%)
FROM_n_of_super_joins__FULL	

Query parameter or value	Descriptive statistics
0	766 (77%)
1	219 (22%)
2	15 (1.5%)
FROM_n_of_joins__INNER	
0	260 (26%)
1	251 (25%)
2	207 (21%)
3	146 (15%)
4	87 (8.7%)
5	35 (3.5%)
6	8 (0.8%)
7	5 (0.5%)
8	1 (0.1%)
FROM_n_of_joins__RIGHT	
0	238 (24%)
1	283 (28%)
2	205 (21%)
3	134 (13%)
4	84 (8.4%)
5	29 (2.9%)
6	15 (1.5%)
7	9 (0.9%)
8	3 (0.3%)
FROM_n_of_processed_rows	[5, 13,620,027,828, 45,900,083,482] 12,857,222,266 / 11,994,133,028
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.81 / 2.28
WHERE_n_of_attribs_of_type__character_varying	
0	463 (46%)
1	307 (31%)
2	152 (15%)
3	59 (5.9%)
4	17 (1.7%)
5	1 (0.1%)
6	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	829 (83%)
1	141 (14%)
2	27 (2.7%)
3	3 (0.3%)
WHERE_n_of_attribs_of_type__integer	
0	452 (45%)
1	318 (32%)
2	154 (15%)
3	56 (5.6%)
4	15 (1.5%)
5	5 (0.5%)

Query parameter or value	Descriptive statistics
WHERE_n_of_attribs_of_type__numeric	
0	631 (63%)
1	271 (27%)
2	72 (7.2%)
3	22 (2.2%)
4	4 (0.4%)
WHERE_n_of_attribs_of_type__character	
0	720 (72%)
1	215 (22%)
2	49 (4.9%)
3	14 (1.4%)
4	2 (0.2%)
WHERE_n_of_pkey_attribs	
0	642 (64%)
1	270 (27%)
2	67 (6.7%)
3	16 (1.6%)
4	5 (0.5%)
WHERE_n_of_connect_OR	[0.00, 1.00, 9.00] 1.37 / 1.75
WHERE_n_of_operators__greater_or_less	
0	416 (42%)
1	299 (30%)
2	160 (16%)
3	85 (8.5%)
4	29 (2.9%)
5	5 (0.5%)
6	5 (0.5%)
7	1 (0.1%)
WHERE_n_of_operators__in	
0	593 (59%)
1	292 (29%)
2	86 (8.6%)
3	24 (2.4%)
4	4 (0.4%)
5	1 (0.1%)
WHERE_n_of_operators__between	
0	617 (62%)
1	267 (27%)
2	96 (9.6%)
3	18 (1.8%)
4	1 (0.1%)
5	1 (0.1%)
WHERE_n_of_operators__like	
0	859 (86%)
1	128 (13%)
2	10 (1.0%)

Query parameter or value	Descriptive statistics
3	3 (0.3%)
WHERE_n_of_non_aggr_func__MONTH	
0	969 (97%)
1	31 (3.1%)
WHERE_n_of_non_aggr_func__SQRT	
0	921 (92%)
1	73 (7.3%)
2	6 (0.6%)
WHERE_n_of_non_aggr_func__TRUNC	
0	967 (97%)
1	33 (3.3%)
WHERE_n_of_non_aggr_func__YEAR	
0	973 (97%)
1	27 (2.7%)
WHERE_n_of_non_aggr_func__LOG	
0	948 (95%)
1	50 (5.0%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__ABS	
0	961 (96%)
1	36 (3.6%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func__DAY	
0	973 (97%)
1	25 (2.5%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__DOW	
0	977 (98%)
1	23 (2.3%)
WHERE_n_of_non_aggr_func__ROUND	
0	973 (97%)
1	25 (2.5%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__FLOOR	
0	973 (97%)
1	27 (2.7%)
WHERE_n_of_all_non_aggr_func	
0	711 (71%)
1	223 (22%)
2	47 (4.7%)
3	14 (1.4%)
4	4 (0.4%)
5	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.34 / 2.09
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.37 / 2.52
HAVING_n_of_main_predicates__non_scalar_subquery	

Query parameter or value	Descriptive statistics
<i>0</i>	434 (43%)
<i>1</i>	209 (21%)
<i>2</i>	178 (18%)
<i>3</i>	179 (18%)
HAVING_n_of_main_predicates__scalar_subquery	
<i>0</i>	444 (44%)
<i>1</i>	167 (17%)
<i>2</i>	212 (21%)
<i>3</i>	177 (18%)
HAVING_n_of_subqueries__non_scalar_subquery	
<i>0</i>	434 (43%)
<i>1</i>	209 (21%)
<i>2</i>	178 (18%)
<i>3</i>	179 (18%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 24.0] 2.3 / 3.5
HAVING_n_of_processed_rows_by_subquery	[0, 18,000,000,000, 826,201,502,046] 41,765,330,808 / 77,498,721,125
ORDER_BY_n_of_columns	
<i>0</i>	298 (30%)
<i>1</i>	401 (40%)
<i>2</i>	206 (21%)
<i>3</i>	75 (7.5%)
<i>4</i>	9 (0.9%)
<i>5</i>	5 (0.5%)
<i>6</i>	5 (0.5%)
<i>7</i>	1 (0.1%)
limit	[1, 491, 1,000] 494 / 289
offset	[0, 33, 997] 259 / 331