

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

Scenario: 1-3jp_ujpp__g__h0-3ss_h0nss

Scale factor: 50 GB

The 1000-query pack was generated on: 2023-11-09

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, 92] 9 / 12
SELECT_n_of_non_aggr_func__ABS	
0	939 (94%)
1	56 (5.6%)
2	5 (0.5%)
SELECT_n_of_non_aggr_func__ROUND	
0	938 (94%)
1	55 (5.5%)
2	6 (0.6%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__RTRIM	
0	716 (72%)
1	211 (21%)
2	54 (5.4%)
3	14 (1.4%)

Query parameter or value	Descriptive statistics
4	4 (0.4%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	708 (71%)
1	220 (22%)
2	52 (5.2%)
3	10 (1.0%)
4	7 (0.7%)
5	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__TRUNC	
0	952 (95%)
1	43 (4.3%)
2	5 (0.5%)
SELECT_n_of_non_aggr_func__UPPER	
0	680 (68%)
1	243 (24%)
2	56 (5.6%)
3	16 (1.6%)
4	3 (0.3%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func__DOW	
0	886 (89%)
1	90 (9.0%)
2	17 (1.7%)
3	3 (0.3%)
4	4 (0.4%)
SELECT_n_of_non_aggr_func__MONTH	
0	844 (84%)
1	125 (13%)
2	17 (1.7%)
3	6 (0.6%)
4	6 (0.6%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__LOWER	
0	697 (70%)
1	231 (23%)
2	57 (5.7%)
3	12 (1.2%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__DAY	
0	911 (91%)
1	79 (7.9%)
2	6 (0.6%)

Query parameter or value	Descriptive statistics
3	4 (0.4%)
SELECT_n_of_non_aggr_func__LOG	
0	885 (89%)
1	101 (10%)
2	11 (1.1%)
3	2 (0.2%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__LTRIM	
0	696 (70%)
1	218 (22%)
2	64 (6.4%)
3	14 (1.4%)
4	5 (0.5%)
5	3 (0.3%)
SELECT_n_of_non_aggr_func__YEAR	
0	864 (86%)
1	105 (11%)
2	22 (2.2%)
3	4 (0.4%)
4	4 (0.4%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__SQRT	
0	841 (84%)
1	130 (13%)
2	17 (1.7%)
3	5 (0.5%)
4	6 (0.6%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__FLOOR	
0	948 (95%)
1	45 (4.5%)
2	7 (0.7%)
SELECT_n_of_all_non_aggr_func	[0.0, 2.0, 43.0] 3.3 / 4.4
SELECT_n_of_aggr_func__COUNT_DISTINCT	
0	737 (74%)
1	94 (9.4%)
2	60 (6.0%)
3	32 (3.2%)
4	31 (3.1%)
5	20 (2.0%)
6	12 (1.2%)
7	11 (1.1%)
8	3 (0.3%)
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 9.00] 0.65 / 1.49
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 10.00] 0.65 / 1.47
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 9.00] 0.68 / 1.52

Query parameter or value	Descriptive statistics
SELECT__n_of__aggr_func__AVG	
0	928 (93%)
1	51 (5.1%)
2	14 (1.4%)
3	7 (0.7%)
SELECT__n_of__aggr_func__SUM	
0	942 (94%)
1	41 (4.1%)
2	13 (1.3%)
3	3 (0.3%)
4	1 (0.1%)
SELECT__n_of__all__aggr_func	[0.00, 2.00, 11.00] 2.87 / 2.31
FROM__n_of__join_paths	
1	340 (34%)
2	343 (34%)
3	317 (32%)
FROM__n_of__super_joins__RIGHT	
0	784 (78%)
1	201 (20%)
2	15 (1.5%)
FROM__n_of__super_joins__FULL	
0	769 (77%)
1	206 (21%)
2	25 (2.5%)
FROM__n_of__super_joins__LEFT	
0	768 (77%)
1	213 (21%)
2	19 (1.9%)
FROM__n_of__joins__INNER	
0	246 (25%)
1	268 (27%)
2	199 (20%)
3	162 (16%)
4	90 (9.0%)
5	21 (2.1%)
6	10 (1.0%)
7	4 (0.4%)
FROM__n_of__joins__RIGHT	
0	246 (25%)
1	288 (29%)
2	204 (20%)
3	128 (13%)
4	77 (7.7%)
5	42 (4.2%)
6	9 (0.9%)
7	4 (0.4%)

Query parameter or value	Descriptive statistics
8	2 (0.2%)
FROM_n_of_processed_rows	[5, 340,241,940, 1,147,549,246] 308,238,294 / 297,261,469
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.79 / 2.20
WHERE_n_of_attribs_of_type__character	
0	744 (74%)
1	210 (21%)
2	39 (3.9%)
3	7 (0.7%)
WHERE_n_of_attribs_of_type__character_varying	
0	419 (42%)
1	330 (33%)
2	171 (17%)
3	62 (6.2%)
4	14 (1.4%)
5	3 (0.3%)
6	1 (0.1%)
WHERE_n_of_attribs_of_type__numeric	
0	632 (63%)
1	257 (26%)
2	86 (8.6%)
3	20 (2.0%)
4	5 (0.5%)
WHERE_n_of_attribs_of_type__date	
0	832 (83%)
1	141 (14%)
2	24 (2.4%)
3	3 (0.3%)
WHERE_n_of_attribs_of_type__integer	
0	474 (47%)
1	303 (30%)
2	153 (15%)
3	50 (5.0%)
4	17 (1.7%)
5	2 (0.2%)
6	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	624 (62%)
1	292 (29%)
2	67 (6.7%)
3	14 (1.4%)
4	3 (0.3%)
WHERE_n_of_connect_OR	
0	465 (47%)
1	181 (18%)
2	134 (13%)

Query parameter or value	Descriptive statistics
3	93 (9.3%)
4	71 (7.1%)
5	31 (3.1%)
6	22 (2.2%)
7	3 (0.3%)
WHERE_n_of_operators__in	
0	600 (60%)
1	285 (29%)
2	83 (8.3%)
3	26 (2.6%)
4	6 (0.6%)
WHERE_n_of_operators__like	
0	843 (84%)
1	143 (14%)
2	13 (1.3%)
3	1 (0.1%)
WHERE_n_of_operators__greater_or_less	
0	432 (43%)
1	297 (30%)
2	162 (16%)
3	72 (7.2%)
4	29 (2.9%)
5	7 (0.7%)
6	1 (0.1%)
WHERE_n_of_operators__between	
0	623 (62%)
1	281 (28%)
2	74 (7.4%)
3	20 (2.0%)
4	2 (0.2%)
WHERE_n_of_non_aggr_func__TRUNC	
0	958 (96%)
1	41 (4.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__SQRT	
0	902 (90%)
1	90 (9.0%)
2	7 (0.7%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func__YEAR	
0	964 (96%)
1	34 (3.4%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__ABS	
0	968 (97%)
1	31 (3.1%)

Query parameter or value	Descriptive statistics
2	1 (0.1%)
WHERE_n_of_non_aggr_func__ROUND	
0	969 (97%)
1	31 (3.1%)
WHERE_n_of_non_aggr_func__DOW	
0	971 (97%)
1	29 (2.9%)
WHERE_n_of_non_aggr_func__LOG	
0	949 (95%)
1	50 (5.0%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DAY	
0	975 (98%)
1	24 (2.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__MONTH	
0	980 (98%)
1	19 (1.9%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__FLOOR	
0	961 (96%)
1	39 (3.9%)
WHERE_n_of_all_non_aggr_func	
0	686 (69%)
1	231 (23%)
2	64 (6.4%)
3	16 (1.6%)
4	3 (0.3%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.38 / 2.07
HAVING_n_of_main_predicates	
0	290 (29%)
1	105 (11%)
2	126 (13%)
3	185 (19%)
4	146 (15%)
5	94 (9.4%)
6	54 (5.4%)
HAVING_n_of_main_predicates__scalar_subquery	
0	433 (43%)
1	200 (20%)
2	162 (16%)
3	205 (21%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 21.0] 2.5 / 3.7
HAVING_n_of_processed_rows_by_subqueries	[0, 6,500,340, 14,500,159,989] 739,910,009 / 1,612,802,254
ORDER_BY_n_of_columns	

Query parameter or value	Descriptive statistics
<i>0</i>	329 (33%)
<i>1</i>	384 (38%)
<i>2</i>	213 (21%)
<i>3</i>	59 (5.9%)
<i>4</i>	8 (0.8%)
<i>5</i>	4 (0.4%)
<i>7</i>	1 (0.1%)
<i>8</i>	2 (0.2%)
limit	[1, 512, 999] 513 / 288
offset	[0, 1, 996] 252 / 321