

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

Scenario: 1-6jp_rjpp__g__h0ss_h0nss

Scale factor: 0.01 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, 9, 155] 15 / 22
SELECT_n_of_non_aggr_func__LOWER	[0.00, 0.00, 11.00] 0.57 / 1.07
SELECT_n_of_non_aggr_func__RTRIM	
0	630 (63%)
1	267 (27%)
10	1 (0.1%)
2	53 (5.3%)
3	26 (2.6%)
4	13 (1.3%)
5	7 (0.7%)
6	3 (0.3%)
SELECT_n_of_non_aggr_func__UPPER	[0.00, 0.00, 11.00] 0.65 / 1.10
SELECT_n_of_non_aggr_func__DAY	
0	892 (89%)
1	84 (8.4%)

Query parameter or value	Descriptive statistics
2	13 (1.3%)
3	4 (0.4%)
4	7 (0.7%)
SELECT_n_of_non_aggr_func__ABS	
0	925 (93%)
1	65 (6.5%)
2	7 (0.7%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func__SUBSTR	[0.00, 0.00, 9.00] 0.60 / 1.04
SELECT_n_of_non_aggr_func__LTRIM	
0	625 (63%)
1	251 (25%)
2	74 (7.4%)
3	26 (2.6%)
4	9 (0.9%)
5	5 (0.5%)
6	5 (0.5%)
7	4 (0.4%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func__DOW	
0	857 (86%)
1	96 (9.6%)
10	1 (0.1%)
2	27 (2.7%)
3	12 (1.2%)
4	4 (0.4%)
5	2 (0.2%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__MONTH	[0.00, 0.00, 10.00] 0.34 / 0.94
SELECT_n_of_non_aggr_func__SQRT	
0	815 (82%)
1	138 (14%)
2	19 (1.9%)
3	18 (1.8%)
4	6 (0.6%)
5	2 (0.2%)
6	1 (0.1%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func__YEAR	
0	834 (83%)
1	114 (11%)
2	27 (2.7%)
3	12 (1.2%)
4	4 (0.4%)
5	4 (0.4%)
6	4 (0.4%)

Query parameter or value	Descriptive statistics
7	1 (0.1%)
SELECT_n_of_non_aggr_func__LOG	
0	878 (88%)
1	95 (9.5%)
2	15 (1.5%)
3	7 (0.7%)
4	2 (0.2%)
5	3 (0.3%)
SELECT_n_of_non_aggr_func__ROUND	
0	917 (92%)
1	66 (6.6%)
2	13 (1.3%)
3	4 (0.4%)
SELECT_n_of_non_aggr_func__TRUNC	
0	915 (92%)
1	72 (7.2%)
2	11 (1.1%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func__FLOOR	
0	922 (92%)
1	69 (6.9%)
2	5 (0.5%)
3	4 (0.4%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 67.0] 4.8 / 7.5
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 11.00] 0.83 / 1.86
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 14.00] 0.88 / 2.06
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 12.00] 0.86 / 1.87
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 12.00] 0.81 / 1.79
SELECT_n_of_aggr_func__SUM	
0	942 (94%)
1	39 (3.9%)
2	13 (1.3%)
3	6 (0.6%)
SELECT_n_of_aggr_func__AVG	
0	927 (93%)
1	52 (5.2%)
2	12 (1.2%)
3	7 (0.7%)
4	2 (0.2%)
SELECT_n_of_all_aggr_func	[0.0, 3.0, 14.0] 3.6 / 3.0
FROM_n_of_join_paths	
1	171 (17%)
2	153 (15%)
3	199 (20%)
4	186 (19%)
5	141 (14%)

Query parameter or value	Descriptive statistics
6	150 (15%)
FROM_n_of_super_joins__LEFT	
0	547 (55%)
1	326 (33%)
2	109 (11%)
3	17 (1.7%)
5	1 (0.1%)
FROM_n_of_super_joins__FULL	
0	551 (55%)
1	303 (30%)
2	119 (12%)
3	25 (2.5%)
4	2 (0.2%)
FROM_n_of_super_joins__RIGHT	
0	568 (57%)
1	301 (30%)
2	114 (11%)
3	14 (1.4%)
4	3 (0.3%)
FROM_n_of_joins__INNER	[0.00, 3.00, 12.00] 3.33 / 2.46
FROM_n_of_joins__RIGHT	[0.00, 3.00, 12.00] 3.29 / 2.43
FROM_n_of_processed_rows	[5, 93,095, 443,305] 118,315 / 98,663
WHERE_n_of_predicates	[0.00, 3.00, 13.00] 3.83 / 2.95
WHERE_n_of_attrbs_of_type__integer	
0	380 (38%)
1	256 (26%)
2	193 (19%)
3	112 (11%)
4	41 (4.1%)
5	14 (1.4%)
6	4 (0.4%)
WHERE_n_of_attrbs_of_type__numeric	
0	559 (56%)
1	293 (29%)
2	107 (11%)
3	27 (2.7%)
4	13 (1.3%)
6	1 (0.1%)
WHERE_n_of_attrbs_of_type__character_varying	
0	363 (36%)
1	298 (30%)
2	172 (17%)
3	104 (10%)
4	41 (4.1%)
5	15 (1.5%)
6	5 (0.5%)

Query parameter or value	Descriptive statistics
7	2 (0.2%)
WHERE_n_of_attribs_of_type__date	
0	803 (80%)
1	161 (16%)
2	31 (3.1%)
3	4 (0.4%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type__character	
0	640 (64%)
1	263 (26%)
2	83 (8.3%)
3	11 (1.1%)
4	3 (0.3%)
WHERE_n_of_pkey_attribs	
0	545 (55%)
1	290 (29%)
2	120 (12%)
3	38 (3.8%)
4	6 (0.6%)
5	1 (0.1%)
WHERE_n_of_connect_OR	[0.00, 1.00, 11.00] 2.23 / 2.41
WHERE_n_of_operators__greater_or_less	
0	315 (32%)
1	287 (29%)
2	191 (19%)
3	109 (11%)
4	65 (6.5%)
5	23 (2.3%)
6	9 (0.9%)
8	1 (0.1%)
WHERE_n_of_operators__in	
0	522 (52%)
1	301 (30%)
2	131 (13%)
3	35 (3.5%)
4	7 (0.7%)
5	3 (0.3%)
6	1 (0.1%)
WHERE_n_of_operators__between	
0	541 (54%)
1	283 (28%)
2	126 (13%)
3	35 (3.5%)
4	8 (0.8%)
5	7 (0.7%)
WHERE_n_of_operators__like	

Query parameter or value	Descriptive statistics
0	790 (79%)
1	174 (17%)
2	34 (3.4%)
3	2 (0.2%)
WHERE_n_of_non_aggr_func__SQRT	
0	884 (88%)
1	107 (11%)
2	9 (0.9%)
WHERE_n_of_non_aggr_func__LOG	
0	927 (93%)
1	70 (7.0%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func__DOW	
0	967 (97%)
1	33 (3.3%)
WHERE_n_of_non_aggr_func__TRUNC	
0	968 (97%)
1	31 (3.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__ABS	
0	961 (96%)
1	38 (3.8%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DAY	
0	969 (97%)
1	31 (3.1%)
WHERE_n_of_non_aggr_func__MONTH	
0	968 (97%)
1	31 (3.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__ROUND	
0	957 (96%)
1	43 (4.3%)
WHERE_n_of_non_aggr_func__FLOOR	
0	955 (96%)
1	45 (4.5%)
WHERE_n_of_non_aggr_func__YEAR	
0	972 (97%)
1	26 (2.6%)
2	2 (0.2%)
WHERE_n_of_all_non_aggr_func	
0	652 (65%)
1	241 (24%)
2	81 (8.1%)
3	20 (2.0%)
4	4 (0.4%)

Query parameter or value	Descriptive statistics
5	2 (0.2%)
GROUP_BY_n_of_columns	[0.00, 3.00, 12.00] 3.24 / 2.92
HAVING_n_of_main_predicates	
0	432 (43%)
1	196 (20%)
2	181 (18%)
3	191 (19%)
ORDER_BY_n_of_columns	
0	285 (29%)
1	341 (34%)
2	247 (25%)
3	96 (9.6%)
4	21 (2.1%)
5	7 (0.7%)
6	2 (0.2%)
9	1 (0.1%)
limit	[1, 515, 998] 513 / 291
offset	[0, 0, 1,000] 250 / 323