

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

Scenario: 1-6jp_rjpp__g__h0ss_h0nss

Scale factor: 100 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, 209] 16 / 24
SELECT_n_of_non_aggr_func__LOWER	
0	598 (60%)
1	271 (27%)
10	1 (0.1%)
15	1 (0.1%)
2	84 (8.4%)
3	22 (2.2%)
4	8 (0.8%)
5	8 (0.8%)
6	7 (0.7%)
SELECT_n_of_non_aggr_func__LTRIM	
0	596 (60%)
1	282 (28%)
2	72 (7.2%)

Query parameter or value	Descriptive statistics
3	23 (2.3%)
4	12 (1.2%)
5	7 (0.7%)
6	5 (0.5%)
7	1 (0.1%)
8	2 (0.2%)
SELECT_n_of_non_aggr_func__ROUND	
0	916 (92%)
1	67 (6.7%)
2	15 (1.5%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func__SUBSTR	[0.00, 0.00, 9.00] 0.64 / 1.08
SELECT_n_of_non_aggr_func__FLOOR	
0	924 (92%)
1	63 (6.3%)
2	9 (0.9%)
3	1 (0.1%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__RTRIM	
0	615 (62%)
1	260 (26%)
11	1 (0.1%)
2	79 (7.9%)
3	17 (1.7%)
4	12 (1.2%)
5	6 (0.6%)
6	8 (0.8%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func__UPPER	[0.00, 0.00, 9.00] 0.62 / 1.05
SELECT_n_of_non_aggr_func__SQRT	
0	795 (80%)
1	138 (14%)
2	29 (2.9%)
3	20 (2.0%)
4	7 (0.7%)
5	7 (0.7%)
6	3 (0.3%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__DAY	
0	889 (89%)
1	87 (8.7%)
2	11 (1.1%)
3	10 (1.0%)
4	1 (0.1%)
5	1 (0.1%)

Query parameter or value	Descriptive statistics
6	1 (0.1%)
SELECT_n_of_non_aggr_func__MONTH	[0.00, 0.00, 16.00] 0.32 / 1.00
SELECT_n_of_non_aggr_func__YEAR	
0	845 (85%)
1	97 (9.7%)
15	1 (0.1%)
2	33 (3.3%)
3	14 (1.4%)
4	6 (0.6%)
5	3 (0.3%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__LOG	
0	856 (86%)
1	97 (9.7%)
2	29 (2.9%)
3	8 (0.8%)
4	6 (0.6%)
5	3 (0.3%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__ABS	
0	923 (92%)
1	62 (6.2%)
2	9 (0.9%)
3	4 (0.4%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__DOW	
0	869 (87%)
1	87 (8.7%)
11	1 (0.1%)
2	23 (2.3%)
3	13 (1.3%)
4	5 (0.5%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func__TRUNC	
0	905 (91%)
1	83 (8.3%)
2	8 (0.8%)
3	2 (0.2%)
4	2 (0.2%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 114.0] 5.1 / 8.5
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 15.00] 0.83 / 1.93
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 12.00] 0.78 / 1.82
SELECT_n_of_aggr_func__AVG	
0	922 (92%)
1	56 (5.6%)
2	15 (1.5%)

Query parameter or value	Descriptive statistics
3	6 (0.6%)
4	1 (0.1%)
SELECT_n_of_aggr_func__SUM	
0	922 (92%)
1	46 (4.6%)
2	28 (2.8%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 13.00] 0.86 / 1.89
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 14.00] 0.94 / 2.02
SELECT_n_of_all_aggr_func	[0.00, 3.00, 15.00] 3.63 / 3.10
FROM_n_of_join_paths	
1	167 (17%)
2	161 (16%)
3	154 (15%)
4	176 (18%)
5	159 (16%)
6	183 (18%)
FROM_n_of_super_joins__FULL	
0	568 (57%)
1	295 (30%)
2	109 (11%)
3	24 (2.4%)
4	4 (0.4%)
FROM_n_of_super_joins__RIGHT	
0	532 (53%)
1	308 (31%)
2	130 (13%)
3	24 (2.4%)
4	6 (0.6%)
FROM_n_of_super_joins__LEFT	
0	545 (55%)
1	309 (31%)
2	114 (11%)
3	26 (2.6%)
4	5 (0.5%)
5	1 (0.1%)
FROM_n_of_joins__INNER	[0.00, 3.00, 14.00] 3.41 / 2.60
FROM_n_of_joins__RIGHT	[0.00, 3.00, 14.00] 3.39 / 2.50
FROM_n_of_processed_rows	[5, 945,038,020, 4,491,227,492] 1,247,422,185 / 1,000,705,984
WHERE_n_of_predicates	[0.00, 3.00, 12.00] 3.84 / 2.96
WHERE_n_of_attribs_of_type__character_varying	
0	369 (37%)
1	293 (29%)
2	163 (16%)

Query parameter or value	Descriptive statistics
3	97 (9.7%)
4	60 (6.0%)
5	9 (0.9%)
6	7 (0.7%)
7	1 (0.1%)
8	1 (0.1%)
WHERE_n_of_attribs_of_type__character	
0	644 (64%)
1	253 (25%)
2	76 (7.6%)
3	19 (1.9%)
4	7 (0.7%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type__integer	
0	365 (37%)
1	304 (30%)
2	173 (17%)
3	85 (8.5%)
4	51 (5.1%)
5	22 (2.2%)
WHERE_n_of_attribs_of_type__numeric	
0	558 (56%)
1	310 (31%)
2	94 (9.4%)
3	31 (3.1%)
4	7 (0.7%)
WHERE_n_of_attribs_of_type__date	
0	788 (79%)
1	171 (17%)
2	35 (3.5%)
3	6 (0.6%)
WHERE_n_of_pkey_attribs	
0	542 (54%)
1	302 (30%)
2	102 (10%)
3	46 (4.6%)
4	4 (0.4%)
5	4 (0.4%)
WHERE_n_of_connect_OR	[0.00, 1.00, 11.00] 2.22 / 2.44
WHERE_n_of_operators__in	
0	526 (53%)
1	286 (29%)
2	137 (14%)
3	41 (4.1%)
4	4 (0.4%)
5	6 (0.6%)

Query parameter or value	Descriptive statistics
WHERE_n_of_operators__between	
0	546 (55%)
1	296 (30%)
2	108 (11%)
3	38 (3.8%)
4	9 (0.9%)
5	2 (0.2%)
6	1 (0.1%)
WHERE_n_of_operators__greater_or_less	
0	320 (32%)
1	273 (27%)
2	186 (19%)
3	113 (11%)
4	64 (6.4%)
5	35 (3.5%)
6	3 (0.3%)
7	6 (0.6%)
WHERE_n_of_operators__like	
0	775 (78%)
1	186 (19%)
2	30 (3.0%)
3	7 (0.7%)
4	2 (0.2%)
WHERE_n_of_non_aggr_func__SQRT	
0	871 (87%)
1	118 (12%)
2	10 (1.0%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func__LOG	
0	927 (93%)
1	70 (7.0%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func__YEAR	
0	963 (96%)
1	37 (3.7%)
WHERE_n_of_non_aggr_func__FLOOR	
0	956 (96%)
1	43 (4.3%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DOW	
0	959 (96%)
1	39 (3.9%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__MONTH	
0	962 (96%)
1	38 (3.8%)

Query parameter or value	Descriptive statistics
WHERE_n_of_non_aggr_func__ROUND	
0	960 (96%)
1	40 (4.0%)
WHERE_n_of_non_aggr_func__DAY	
0	961 (96%)
1	39 (3.9%)
WHERE_n_of_non_aggr_func__ABS	
0	957 (96%)
1	42 (4.2%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__TRUNC	
0	971 (97%)
1	28 (2.8%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	610 (61%)
1	280 (28%)
2	83 (8.3%)
3	21 (2.1%)
4	6 (0.6%)
GROUP_BY_n_of_columns	[0.00, 3.00, 12.00] 3.38 / 2.95
HAVING_n_of_main_predicates	
0	411 (41%)
1	196 (20%)
2	195 (20%)
3	198 (20%)
ORDER_BY_n_of_columns	[0.00, 1.00, 9.00] 1.31 / 1.23
limit	[1, 492, 1,000] 495 / 289
offset	[0, 19, 999] 246 / 319