

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

Scenario: 1-6jp_rjpp__g__h0ss_h0nss

Scale factor: 1 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, 195] 16 / 25
SELECT_n_of_non_aggr_func__LOG	
0	847 (85%)
1	114 (11%)
2	19 (1.9%)
3	11 (1.1%)
4	4 (0.4%)
5	3 (0.3%)
6	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__LOWER	[0.00, 0.00, 10.00] 0.60 / 1.04
SELECT_n_of_non_aggr_func__LTRIM	[0.00, 0.00, 12.00] 0.64 / 1.20
SELECT_n_of_non_aggr_func__SUBSTR	[0.00, 0.00, 10.00] 0.63 / 1.19
SELECT_n_of_non_aggr_func__UPPER	
0	594 (59%)

Query parameter or value	Descriptive statistics
1	281 (28%)
10	1 (0.1%)
2	69 (6.9%)
3	29 (2.9%)
4	15 (1.5%)
5	7 (0.7%)
6	3 (0.3%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__RTRIM	
0	596 (60%)
1	247 (25%)
2	100 (10%)
3	26 (2.6%)
4	15 (1.5%)
5	5 (0.5%)
6	6 (0.6%)
7	4 (0.4%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__ROUND	
0	914 (91%)
1	74 (7.4%)
2	9 (0.9%)
3	2 (0.2%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__SQRT	
0	789 (79%)
1	154 (15%)
2	21 (2.1%)
3	13 (1.3%)
4	9 (0.9%)
5	4 (0.4%)
6	2 (0.2%)
7	5 (0.5%)
8	3 (0.3%)
SELECT_n_of_non_aggr_func__DOW	
0	844 (84%)
1	112 (11%)
2	24 (2.4%)
3	9 (0.9%)
4	7 (0.7%)
5	2 (0.2%)
6	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__FLOOR	
0	903 (90%)
1	80 (8.0%)

Query parameter or value	Descriptive statistics
2	15 (1.5%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__TRUNC	
0	917 (92%)
1	63 (6.3%)
2	15 (1.5%)
3	4 (0.4%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__YEAR	
0	820 (82%)
1	115 (12%)
10	1 (0.1%)
2	32 (3.2%)
3	15 (1.5%)
4	10 (1.0%)
5	4 (0.4%)
6	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__MONTH	[0.00, 0.00, 10.00] 0.38 / 1.03
SELECT_n_of_non_aggr_func__ABS	
0	912 (91%)
1	75 (7.5%)
2	10 (1.0%)
3	1 (0.1%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__DAY	
0	872 (87%)
1	101 (10%)
2	15 (1.5%)
3	5 (0.5%)
4	4 (0.4%)
6	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0, 3, 75] 5 / 9
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 12.00] 0.88 / 2.02
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 14.00] 0.80 / 1.95
SELECT_n_of_aggr_func__SUM	
0	948 (95%)
1	35 (3.5%)
2	13 (1.3%)
3	2 (0.2%)
4	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 14.00] 0.85 / 2.04
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 9.00] 0.86 / 1.78

Query parameter or value	Descriptive statistics
SELECT__n_of__aggr__func__AVG	
0	931 (93%)
1	42 (4.2%)
2	19 (1.9%)
3	7 (0.7%)
4	1 (0.1%)
SELECT__n_of__all__aggr__func	[0.0, 3.0, 14.0] 3.6 / 3.2
FROM__n_of__join__paths	
1	161 (16%)
2	177 (18%)
3	163 (16%)
4	163 (16%)
5	176 (18%)
6	160 (16%)
FROM__n_of__super__joins__FULL	
0	546 (55%)
1	329 (33%)
2	94 (9.4%)
3	29 (2.9%)
4	2 (0.2%)
FROM__n_of__super__joins__LEFT	
0	548 (55%)
1	320 (32%)
2	107 (11%)
3	23 (2.3%)
4	1 (0.1%)
5	1 (0.1%)
FROM__n_of__super__joins__RIGHT	
0	550 (55%)
1	301 (30%)
2	116 (12%)
3	29 (2.9%)
4	3 (0.3%)
5	1 (0.1%)
FROM__n_of__joins__INNER	[0.00, 3.00, 12.00] 3.29 / 2.50
FROM__n_of__joins__RIGHT	[0.00, 3.00, 15.00] 3.49 / 2.54
FROM__n_of__processed__rows	[5, 9,405,689, 44,913,661] 12,066,898 / 9,757,362
WHERE__n_of__predicates	[0.00, 3.00, 13.00] 3.74 / 2.98
WHERE__n_of__attribs__of__type__character__varying	
0	371 (37%)
1	294 (29%)
2	188 (19%)
3	79 (7.9%)
4	41 (4.1%)
5	21 (2.1%)
6	6 (0.6%)

Query parameter or value	Descriptive statistics
WHERE_n_of_attribs_of_type__integer	
0	386 (39%)
1	277 (28%)
2	193 (19%)
3	86 (8.6%)
4	41 (4.1%)
5	11 (1.1%)
6	6 (0.6%)
WHERE_n_of_attribs_of_type__numeric	
0	596 (60%)
1	260 (26%)
2	107 (11%)
3	27 (2.7%)
4	5 (0.5%)
5	5 (0.5%)
WHERE_n_of_attribs_of_type__character	
0	655 (66%)
1	243 (24%)
2	79 (7.9%)
3	20 (2.0%)
4	3 (0.3%)
WHERE_n_of_attribs_of_type__date	
0	772 (77%)
1	183 (18%)
2	37 (3.7%)
3	8 (0.8%)
WHERE_n_of_pkey_attribs	
0	541 (54%)
1	296 (30%)
2	122 (12%)
3	32 (3.2%)
4	9 (0.9%)
WHERE_n_of_connect_OR	[0.00, 1.00, 12.00] 2.16 / 2.45
WHERE_n_of_operators__in	
0	554 (55%)
1	284 (28%)
2	108 (11%)
3	41 (4.1%)
4	12 (1.2%)
5	1 (0.1%)
WHERE_n_of_operators__like	
0	802 (80%)
1	168 (17%)
2	28 (2.8%)
3	2 (0.2%)
WHERE_n_of_operators__greater_or_less	

Query parameter or value	Descriptive statistics
0	338 (34%)
1	283 (28%)
2	178 (18%)
3	99 (9.9%)
4	58 (5.8%)
5	25 (2.5%)
6	16 (1.6%)
7	3 (0.3%)
WHERE_n_of_operators__between	
0	547 (55%)
1	275 (28%)
2	120 (12%)
3	43 (4.3%)
4	11 (1.1%)
5	2 (0.2%)
6	1 (0.1%)
7	1 (0.1%)
WHERE_n_of_non_aggr_func__ABS	
0	957 (96%)
1	43 (4.3%)
WHERE_n_of_non_aggr_func__MONTH	
0	950 (95%)
1	49 (4.9%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__TRUNC	
0	965 (97%)
1	34 (3.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__SQRT	
0	862 (86%)
1	125 (13%)
2	13 (1.3%)
WHERE_n_of_non_aggr_func__ROUND	
0	968 (97%)
1	31 (3.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DAY	
0	966 (97%)
1	34 (3.4%)
WHERE_n_of_non_aggr_func__YEAR	
0	960 (96%)
1	38 (3.8%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__DOW	
0	968 (97%)
1	30 (3.0%)

Query parameter or value	Descriptive statistics
2	2 (0.2%)
WHERE_n_of_non_aggr_func__LOG	
0	929 (93%)
1	64 (6.4%)
2	7 (0.7%)
WHERE_n_of_non_aggr_func__FLOOR	
0	957 (96%)
1	39 (3.9%)
2	4 (0.4%)
WHERE_n_of_all_non_aggr_func	
0	625 (63%)
1	249 (25%)
2	84 (8.4%)
3	36 (3.6%)
4	6 (0.6%)
GROUP_BY_n_of_columns	[0.00, 3.00, 12.00] 3.17 / 2.88
HAVING_n_of_main_predicates	
0	450 (45%)
1	186 (19%)
2	170 (17%)
3	194 (19%)
ORDER_BY_n_of_columns	
0	307 (31%)
1	310 (31%)
2	249 (25%)
3	100 (10%)
4	19 (1.9%)
5	7 (0.7%)
6	8 (0.8%)
limit	[2, 518, 1,000] 508 / 289
offset	[0, 0, 1,000] 246 / 324