

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

Scenario: 1-3jp_ujpp__g__h0ss_h0nss

Scale factor: 2000 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, 113] 10 / 13
SELECT_n_of_non_aggr_func__LOWER	
0	700 (70%)
1	230 (23%)
2	58 (5.8%)
3	11 (1.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__SQRT	
0	849 (85%)
1	126 (13%)
2	18 (1.8%)
3	4 (0.4%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__UPPER	

Query parameter or value	Descriptive statistics
0	713 (71%)
1	235 (24%)
2	37 (3.7%)
3	10 (1.0%)
4	3 (0.3%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___DOW	
0	901 (90%)
1	78 (7.8%)
2	16 (1.6%)
3	3 (0.3%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func___FLOOR	
0	934 (93%)
1	59 (5.9%)
2	7 (0.7%)
SELECT_n_of_non_aggr_func___LTRIM	
0	719 (72%)
1	209 (21%)
2	51 (5.1%)
3	12 (1.2%)
4	7 (0.7%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___MONTH	
0	854 (85%)
1	106 (11%)
2	27 (2.7%)
3	10 (1.0%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___ROUND	
0	948 (95%)
1	48 (4.8%)
2	4 (0.4%)
SELECT_n_of_non_aggr_func___RTRIM	
0	695 (70%)
1	232 (23%)
2	52 (5.2%)
3	15 (1.5%)
4	3 (0.3%)
5	3 (0.3%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	698 (70%)
1	223 (22%)
2	57 (5.7%)

Query parameter or value	Descriptive statistics
3	13 (1.3%)
4	8 (0.8%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__DAY	
0	920 (92%)
1	70 (7.0%)
2	9 (0.9%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__YEAR	
0	880 (88%)
1	94 (9.4%)
2	16 (1.6%)
3	6 (0.6%)
4	3 (0.3%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__LOG	
0	879 (88%)
1	100 (10%)
2	14 (1.4%)
3	4 (0.4%)
4	1 (0.1%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func__TRUNC	
0	946 (95%)
1	45 (4.5%)
2	8 (0.8%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__ABS	
0	938 (94%)
1	55 (5.5%)
2	6 (0.6%)
3	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.0, 2.0, 39.0] 3.1 / 4.4
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 9.00] 0.57 / 1.32
SELECT_n_of_aggr_func__COUNT_DISTINCT	
0	765 (77%)
1	78 (7.8%)
2	61 (6.1%)
3	38 (3.8%)
4	23 (2.3%)
5	14 (1.4%)
6	12 (1.2%)
7	6 (0.6%)
8	3 (0.3%)
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 10.00] 0.75 / 1.60
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 9.00] 0.53 / 1.25

Query parameter or value	Descriptive statistics
SELECT__n_of__aggr_func__SUM	
0	942 (94%)
1	45 (4.5%)
2	11 (1.1%)
3	1 (0.1%)
4	1 (0.1%)
SELECT__n_of__aggr_func__AVG	
0	939 (94%)
1	49 (4.9%)
2	11 (1.1%)
3	1 (0.1%)
SELECT__n_of__all__aggr_func	[0.00, 2.00, 11.00] 2.61 / 2.24
FROM__n_of__join_paths	
1	338 (34%)
2	339 (34%)
3	323 (32%)
FROM__n_of__super__joins__FULL	
0	782 (78%)
1	197 (20%)
2	21 (2.1%)
FROM__n_of__super__joins__LEFT	
0	776 (78%)
1	203 (20%)
2	21 (2.1%)
FROM__n_of__super__joins__RIGHT	
0	775 (78%)
1	203 (20%)
2	22 (2.2%)
FROM__n_of__joins__INNER	
0	225 (23%)
1	244 (24%)
2	250 (25%)
3	130 (13%)
4	84 (8.4%)
5	44 (4.4%)
6	16 (1.6%)
7	6 (0.6%)
8	1 (0.1%)
FROM__n_of__joins__RIGHT	
0	247 (25%)
1	264 (26%)
2	221 (22%)
3	142 (14%)
4	81 (8.1%)
5	34 (3.4%)
6	9 (0.9%)

Query parameter or value	Descriptive statistics
7	2 (0.2%)
FROM_n_of_processed_rows	[5, 13,620,027,831, 45,900,083,457] 13,134,742,849 / 12,082,516,681
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.82 / 2.30
WHERE_n_of_attribs_of_type__character_varying	
0	453 (45%)
1	278 (28%)
2	176 (18%)
3	69 (6.9%)
4	20 (2.0%)
5	3 (0.3%)
6	1 (0.1%)
WHERE_n_of_attribs_of_type__integer	
0	440 (44%)
1	343 (34%)
2	144 (14%)
3	44 (4.4%)
4	27 (2.7%)
5	2 (0.2%)
WHERE_n_of_attribs_of_type__numeric	
0	647 (65%)
1	253 (25%)
2	80 (8.0%)
3	16 (1.6%)
4	4 (0.4%)
WHERE_n_of_attribs_of_type__character	
0	737 (74%)
1	197 (20%)
2	54 (5.4%)
3	11 (1.1%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	848 (85%)
1	121 (12%)
2	31 (3.1%)
WHERE_n_of_pkey_attribs	
0	606 (61%)
1	285 (29%)
2	83 (8.3%)
3	20 (2.0%)
4	6 (0.6%)
WHERE_n_of_connect_OR	
0	442 (44%)
1	199 (20%)
2	120 (12%)
3	100 (10%)

Query parameter or value	Descriptive statistics
4	58 (5.8%)
5	42 (4.2%)
6	22 (2.2%)
7	13 (1.3%)
8	4 (0.4%)
WHERE_n_of_operators__in	
0	621 (62%)
1	274 (27%)
2	80 (8.0%)
3	23 (2.3%)
4	2 (0.2%)
WHERE_n_of_operators__between	
0	605 (61%)
1	265 (27%)
2	102 (10%)
3	24 (2.4%)
4	4 (0.4%)
WHERE_n_of_operators__greater_or_less	
0	408 (41%)
1	317 (32%)
2	157 (16%)
3	80 (8.0%)
4	26 (2.6%)
5	8 (0.8%)
6	2 (0.2%)
7	2 (0.2%)
WHERE_n_of_operators__like	
0	821 (82%)
1	159 (16%)
2	20 (2.0%)
WHERE_n_of_non_aggr_func__FLOOR	
0	961 (96%)
1	37 (3.7%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__SQRT	
0	897 (90%)
1	96 (9.6%)
2	7 (0.7%)
WHERE_n_of_non_aggr_func__DAY	
0	981 (98%)
1	19 (1.9%)
WHERE_n_of_non_aggr_func__DOW	
0	976 (98%)
1	24 (2.4%)
WHERE_n_of_non_aggr_func__TRUNC	
0	968 (97%)

Query parameter or value	Descriptive statistics
1	32 (3.2%)
WHERE_n_of_non_aggr_func__YEAR	
0	974 (97%)
1	26 (2.6%)
WHERE_n_of_non_aggr_func__LOG	
0	948 (95%)
1	49 (4.9%)
2	3 (0.3%)
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	980 (98%)
1	20 (2.0%)
WHERE_n_of_non_aggr_func__ABS	
0	972 (97%)
1	27 (2.7%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	704 (70%)
1	223 (22%)
2	56 (5.6%)
3	14 (1.4%)
4	3 (0.3%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.34 / 2.11
HAVING_n_of_main_predicates	
0	426 (43%)
1	179 (18%)
2	195 (20%)
3	200 (20%)
ORDER_BY_n_of_columns	
0	268 (27%)
1	400 (40%)
2	241 (24%)
3	66 (6.6%)
4	15 (1.5%)
5	7 (0.7%)
6	1 (0.1%)
7	2 (0.2%)
limit	[1, 487, 1,000] 495 / 293
offset	[0, 0, 998] 244 / 324