

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

Scenario: 1-3jp_rjpp__g__h0ss_h0nss

Scale factor: 5 GB

The 1000-query pack was generated on: 2023-11-10

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, 91] 10 / 13
SELECT_n_of_non_aggr_func__LOG	
0	918 (92%)
1	69 (6.9%)
2	12 (1.2%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__LOWER	
0	664 (66%)
1	252 (25%)
2	57 (5.7%)
3	20 (2.0%)
4	6 (0.6%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__UPPER	
0	693 (69%)

Query parameter or value	Descriptive statistics
1	239 (24%)
2	50 (5.0%)
3	11 (1.1%)
4	7 (0.7%)
SELECT_n_of_non_aggr_func__DAY	
0	909 (91%)
1	79 (7.9%)
2	9 (0.9%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func__MONTH	
0	830 (83%)
1	131 (13%)
2	22 (2.2%)
3	11 (1.1%)
4	4 (0.4%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func__RTRIM	
0	656 (66%)
1	259 (26%)
2	57 (5.7%)
3	19 (1.9%)
4	5 (0.5%)
5	4 (0.4%)
SELECT_n_of_non_aggr_func__TRUNC	
0	936 (94%)
1	63 (6.3%)
2	1 (0.1%)
SELECT_n_of_non_aggr_func__LTRIM	
0	705 (71%)
1	221 (22%)
2	53 (5.3%)
3	14 (1.4%)
4	3 (0.3%)
5	4 (0.4%)
SELECT_n_of_non_aggr_func__ABS	
0	932 (93%)
1	62 (6.2%)
2	5 (0.5%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__DOW	
0	875 (88%)
1	97 (9.7%)
2	21 (2.1%)
3	3 (0.3%)
4	3 (0.3%)
7	1 (0.1%)

Query parameter or value	Descriptive statistics
SELECT__n_of__non__aggr__func___SQRT	
0	848 (85%)
1	115 (12%)
2	24 (2.4%)
3	9 (0.9%)
4	3 (0.3%)
5	1 (0.1%)
SELECT__n_of__non__aggr__func___SUBSTR	
0	684 (68%)
1	241 (24%)
2	52 (5.2%)
3	13 (1.3%)
4	4 (0.4%)
5	3 (0.3%)
6	3 (0.3%)
SELECT__n_of__non__aggr__func___YEAR	
0	866 (87%)
1	103 (10%)
2	20 (2.0%)
3	6 (0.6%)
4	4 (0.4%)
5	1 (0.1%)
SELECT__n_of__non__aggr__func___FLOOR	
0	929 (93%)
1	64 (6.4%)
2	5 (0.5%)
3	2 (0.2%)
SELECT__n_of__non__aggr__func___ROUND	
0	940 (94%)
1	55 (5.5%)
2	4 (0.4%)
3	1 (0.1%)
SELECT__n_of__all__non__aggr__func	[0.0, 2.0, 42.0] 3.4 / 4.6
SELECT__n_of__aggr__func___MIN	[0.00, 0.00, 9.00] 0.74 / 1.54
SELECT__n_of__aggr__func___COUNT_DISTINCT	
0	756 (76%)
1	78 (7.8%)
2	56 (5.6%)
3	46 (4.6%)
4	22 (2.2%)
5	23 (2.3%)
6	9 (0.9%)
7	6 (0.6%)
8	4 (0.4%)
SELECT__n_of__aggr__func___MAX	[0.00, 0.00, 9.00] 0.61 / 1.41
SELECT__n_of__aggr__func___AVG	

Query parameter or value	Descriptive statistics
0	945 (95%)
1	38 (3.8%)
2	13 (1.3%)
3	2 (0.2%)
4	2 (0.2%)
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 10.00] 0.62 / 1.44
SELECT_n_of_aggr_func__SUM	
0	950 (95%)
1	34 (3.4%)
2	11 (1.1%)
3	3 (0.3%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 10.00] 2.78 / 2.30
FROM_n_of_join_paths	
1	347 (35%)
2	326 (33%)
3	327 (33%)
FROM_n_of_super_joins__LEFT	
0	770 (77%)
1	208 (21%)
2	22 (2.2%)
FROM_n_of_super_joins__FULL	
0	755 (76%)
1	223 (22%)
2	22 (2.2%)
FROM_n_of_super_joins__RIGHT	
0	795 (80%)
1	187 (19%)
2	18 (1.8%)
FROM_n_of_joins__INNER	
0	232 (23%)
1	254 (25%)
2	220 (22%)
3	154 (15%)
4	81 (8.1%)
5	44 (4.4%)
6	10 (1.0%)
7	4 (0.4%)
8	1 (0.1%)
FROM_n_of_joins__RIGHT	
0	237 (24%)
1	291 (29%)
2	201 (20%)
3	126 (13%)
4	88 (8.8%)

Query parameter or value	Descriptive statistics
5	34 (3.4%)
6	17 (1.7%)
7	4 (0.4%)
8	2 (0.2%)
FROM_n_of_processed_rows	[5, 34,050,188, 114,749,827] 33,800,719 / 30,503,330
WHERE_n_of_predicates	[0.00, 2.00, 9.00] 2.71 / 2.24
WHERE_n_of_attribs_of_type__numeric	
0	670 (67%)
1	229 (23%)
2	85 (8.5%)
3	14 (1.4%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type__character_varying	
0	472 (47%)
1	290 (29%)
2	154 (15%)
3	66 (6.6%)
4	13 (1.3%)
5	3 (0.3%)
6	2 (0.2%)
WHERE_n_of_attribs_of_type__integer	
0	476 (48%)
1	306 (31%)
2	146 (15%)
3	51 (5.1%)
4	16 (1.6%)
5	5 (0.5%)
WHERE_n_of_attribs_of_type__character	
0	717 (72%)
1	220 (22%)
2	51 (5.1%)
3	10 (1.0%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type__date	
0	843 (84%)
1	134 (13%)
2	20 (2.0%)
3	3 (0.3%)
WHERE_n_of_pkey_attribs	
0	650 (65%)
1	264 (26%)
2	65 (6.5%)
3	20 (2.0%)
4	1 (0.1%)
WHERE_n_of_connect_OR	
0	468 (47%)

Query parameter or value	Descriptive statistics
1	176 (18%)
2	126 (13%)
3	104 (10%)
4	58 (5.8%)
5	42 (4.2%)
6	16 (1.6%)
7	10 (1.0%)
WHERE_n_of_operators__in	
0	613 (61%)
1	279 (28%)
2	90 (9.0%)
3	13 (1.3%)
4	5 (0.5%)
WHERE_n_of_operators__between	
0	600 (60%)
1	283 (28%)
2	85 (8.5%)
3	26 (2.6%)
4	5 (0.5%)
5	1 (0.1%)
WHERE_n_of_operators__greater_or_less	
0	432 (43%)
1	290 (29%)
2	165 (17%)
3	76 (7.6%)
4	28 (2.8%)
5	6 (0.6%)
6	2 (0.2%)
7	1 (0.1%)
WHERE_n_of_operators__like	
0	848 (85%)
1	133 (13%)
2	18 (1.8%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func__FLOOR	
0	975 (98%)
1	25 (2.5%)
WHERE_n_of_non_aggr_func__SQRT	
0	906 (91%)
1	93 (9.3%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DAY	
0	974 (97%)
1	26 (2.6%)
WHERE_n_of_non_aggr_func__ROUND	
0	976 (98%)

Query parameter or value	Descriptive statistics
1	24 (2.4%)
WHERE_n_of_non_aggr_func__MONTH	
0	980 (98%)
1	19 (1.9%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__ABS	
0	968 (97%)
1	30 (3.0%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__YEAR	
0	983 (98%)
1	17 (1.7%)
WHERE_n_of_non_aggr_func__LOG	
0	948 (95%)
1	50 (5.0%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__TRUNC	
0	969 (97%)
1	30 (3.0%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DOW	
0	967 (97%)
1	33 (3.3%)
WHERE_n_of_all_non_aggr_func	
0	728 (73%)
1	197 (20%)
2	61 (6.1%)
3	14 (1.4%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.44 / 2.10
HAVING_n_of_main_predicates	
0	431 (43%)
1	197 (20%)
2	189 (19%)
3	183 (18%)
ORDER_BY_n_of_columns	
0	323 (32%)
1	384 (38%)
2	213 (21%)
3	59 (5.9%)
4	14 (1.4%)
5	4 (0.4%)
6	3 (0.3%)
limit	[1, 491, 999] 501 / 289
offset	[0, 42, 1,000] 263 / 322