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

Scenario: 1-3jp_rjjp__g__h0-3ss_h0nss

Scale factor: 2 GB

The 1000-query pack was generated on: 2023-11-06

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, 85] 10 / 12
SELECT_n_of_non_aggr_func___LOWER	
0	681 (68%)
1	247 (25%)
2	56 (5.6%)
3	9 (0.9%)
4	5 (0.5%)
6	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___LTRIM	
0	697 (70%)
1	231 (23%)
2	52 (5.2%)
3	16 (1.6%)
4	3 (0.3%)

Query parameter or value	Descriptive statistics
5	1 (0.1%)
SELECT_n_of_non_aggr_func___UPPER	
0	669 (67%)
1	243 (24%)
2	62 (6.2%)
3	18 (1.8%)
4	5 (0.5%)
5	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___DAY	
0	910 (91%)
1	73 (7.3%)
2	13 (1.3%)
3	4 (0.4%)
SELECT_n_of_non_aggr_func___MONTH	
0	862 (86%)
1	106 (11%)
2	21 (2.1%)
3	5 (0.5%)
4	5 (0.5%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___LOG	
0	888 (89%)
1	86 (8.6%)
2	18 (1.8%)
3	6 (0.6%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___TRUNC	
0	926 (93%)
1	69 (6.9%)
2	5 (0.5%)
SELECT_n_of_non_aggr_func___RTRIM	
0	715 (72%)
1	217 (22%)
2	53 (5.3%)
3	12 (1.2%)
4	3 (0.3%)
SELECT_n_of_non_aggr_func___DOW	
0	893 (89%)
1	90 (9.0%)
2	11 (1.1%)
3	6 (0.6%)
SELECT_n_of_non_aggr_func___SQRT	
0	845 (85%)
1	129 (13%)

Query parameter or value	Descriptive statistics
2	18 (1.8%)
3	5 (0.5%)
4	1 (0.1%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	685 (69%)
1	242 (24%)
2	51 (5.1%)
3	18 (1.8%)
4	2 (0.2%)
5	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___YEAR	
0	872 (87%)
1	105 (11%)
2	16 (1.6%)
3	3 (0.3%)
4	2 (0.2%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___ROUND	
0	936 (94%)
1	61 (6.1%)
2	2 (0.2%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___ABS	
0	943 (94%)
1	47 (4.7%)
2	9 (0.9%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___FLOOR	
0	947 (95%)
1	51 (5.1%)
2	2 (0.2%)
SELECT_n_of_all_non_aggr_func	[0.0, 2.0, 37.0] 3.3 / 4.3
SELECT_n_of_aggr_func___COUNT_DISTINCT	
0	740 (74%)
1	111 (11%)
2	57 (5.7%)
3	41 (4.1%)
4	24 (2.4%)
5	14 (1.4%)
6	6 (0.6%)
7	3 (0.3%)
8	4 (0.4%)
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 9.00] 0.58 / 1.35

Query parameter or value	Descriptive statistics
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 11.00] 0.62 / 1.41
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 9.00] 0.70 / 1.52
SELECT_n_of_aggr_func___SUM	
0	939 (94%)
1	40 (4.0%)
2	17 (1.7%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_aggr_func___AVG	
0	936 (94%)
1	44 (4.4%)
2	15 (1.5%)
3	4 (0.4%)
4	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 12.00] 2.68 / 2.23
FROM_n_of_join_paths	
1	326 (33%)
2	351 (35%)
3	323 (32%)
FROM_n_of_super_joins___RIGHT	
0	764 (76%)
1	212 (21%)
2	24 (2.4%)
FROM_n_of_super_joins___FULL	
0	795 (80%)
1	194 (19%)
2	11 (1.1%)
FROM_n_of_super_joins___LEFT	
0	759 (76%)
1	223 (22%)
2	18 (1.8%)
FROM_n_of_joins___INNER	
0	227 (23%)
1	246 (25%)
2	222 (22%)
3	175 (18%)
4	84 (8.4%)
5	28 (2.8%)
6	15 (1.5%)
7	2 (0.2%)
8	1 (0.1%)
FROM_n_of_joins___RIGHT	
0	251 (25%)
1	254 (25%)
2	199 (20%)
3	162 (16%)

Query parameter or value	Descriptive statistics
4	82 (8.2%)
5	23 (2.3%)
6	23 (2.3%)
7	5 (0.5%)
8	1 (0.1%)
FROM_n_of_processed_rows	[5, 13,615,801, 45,887,206] 12,970,850 / 11,658,818
WHERE_n_of_predicates	[0.00, 2.00, 9.00] 2.81 / 2.18
WHERE_n_of_attribs_of_type___date	
0	846 (85%)
1	132 (13%)
2	19 (1.9%)
3	3 (0.3%)
WHERE_n_of_attribs_of_type___character_varying	
0	424 (42%)
1	343 (34%)
2	146 (15%)
3	68 (6.8%)
4	15 (1.5%)
5	4 (0.4%)
WHERE_n_of_attribs_of_type___integer	
0	452 (45%)
1	330 (33%)
2	148 (15%)
3	52 (5.2%)
4	16 (1.6%)
5	2 (0.2%)
WHERE_n_of_attribs_of_type___numeric	
0	636 (64%)
1	267 (27%)
2	73 (7.3%)
3	21 (2.1%)
4	3 (0.3%)
WHERE_n_of_attribs_of_type___character	
0	719 (72%)
1	211 (21%)
2	59 (5.9%)
3	10 (1.0%)
4	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	620 (62%)
1	288 (29%)
2	74 (7.4%)
3	15 (1.5%)
4	3 (0.3%)
WHERE_n_of_connect_OR	
0	449 (45%)

Query parameter or value	Descriptive statistics
1	182 (18%)
2	135 (14%)
3	106 (11%)
4	63 (6.3%)
5	35 (3.5%)
6	21 (2.1%)
7	7 (0.7%)
8	2 (0.2%)
WHERE_n_of_operators___greater_or_less	
0	415 (42%)
1	295 (30%)
2	180 (18%)
3	76 (7.6%)
4	30 (3.0%)
5	2 (0.2%)
6	1 (0.1%)
7	1 (0.1%)
WHERE_n_of_operators___in	
0	612 (61%)
1	266 (27%)
2	101 (10%)
3	17 (1.7%)
4	3 (0.3%)
5	1 (0.1%)
WHERE_n_of_operators___between	
0	627 (63%)
1	258 (26%)
2	94 (9.4%)
3	19 (1.9%)
4	1 (0.1%)
5	1 (0.1%)
WHERE_n_of_operators___like	
0	839 (84%)
1	140 (14%)
2	21 (2.1%)
WHERE_n_of_non_aggr_func___LOG	
0	951 (95%)
1	48 (4.8%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___SQRT	
0	906 (91%)
1	89 (8.9%)
2	5 (0.5%)
WHERE_n_of_non_aggr_func___TRUNC	
0	966 (97%)
1	34 (3.4%)

Query parameter or value	Descriptive statistics
WHERE_n_of_non_aggr_func___DAY	
0	980 (98%)
1	19 (1.9%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___YEAR	
0	972 (97%)
1	26 (2.6%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___MONTH	
0	967 (97%)
1	33 (3.3%)
WHERE_n_of_non_aggr_func___ROUND	
0	965 (97%)
1	35 (3.5%)
WHERE_n_of_non_aggr_func___ABS	
0	972 (97%)
1	26 (2.6%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___DOW	
0	978 (98%)
1	22 (2.2%)
WHERE_n_of_non_aggr_func___FLOOR	
0	967 (97%)
1	33 (3.3%)
WHERE_n_of_all_non_aggr_func	
0	705 (71%)
1	222 (22%)
2	58 (5.8%)
3	11 (1.1%)
4	4 (0.4%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.33 / 2.04
HAVING_n_of_main_predicates	
0	282 (28%)
1	103 (10%)
2	143 (14%)
3	195 (20%)
4	133 (13%)
5	97 (9.7%)
6	47 (4.7%)
HAVING_n_of_main_predicates___scalar_subquery	
0	422 (42%)
1	212 (21%)
2	192 (19%)
3	174 (17%)
HAVING_n_of_subqueries___scalar_subquery	[0.0, 1.0, 21.0] 2.3 / 3.5
HAVING_n_of_processed_rows_by_subqueries	[0, 1,610,083, 803,459,460] 29,175,302 / 68,674,720

Query parameter or value	Descriptive statistics
ORDER_BY_n_of_columns	
0	293 (29%)
1	366 (37%)
2	255 (26%)
3	63 (6.3%)
4	14 (1.4%)
5	5 (0.5%)
6	3 (0.3%)
7	1 (0.1%)
limit	[1, 488, 999] 494 / 288
offset	[0, 0, 995] 239 / 315