

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

```
## Scenario: 1-3jp_ujpp__g__h0-3ss_h0nss

## Scale factor: 200 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, 110] 9 / 12
SELECT_n_of_non_aggr_func__DAY	
0	914 (91%)
1	70 (7.0%)
2	8 (0.8%)
3	6 (0.6%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__RTRIM	
0	670 (67%)
1	259 (26%)
2	59 (5.9%)
3	8 (0.8%)
4	3 (0.3%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__UPPER	

Query parameter or value	Descriptive statistics
0	680 (68%)
1	255 (26%)
2	49 (4.9%)
3	12 (1.2%)
4	1 (0.1%)
5	2 (0.2%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___SQRT	
0	846 (85%)
1	136 (14%)
2	11 (1.1%)
3	4 (0.4%)
4	2 (0.2%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___LTRIM	
0	676 (68%)
1	253 (25%)
2	59 (5.9%)
3	7 (0.7%)
4	2 (0.2%)
6	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___LOWER	
0	713 (71%)
1	213 (21%)
2	58 (5.8%)
3	14 (1.4%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	692 (69%)
1	249 (25%)
2	41 (4.1%)
3	10 (1.0%)
4	5 (0.5%)
5	3 (0.3%)
SELECT_n_of_non_aggr_func___FLOOR	
0	952 (95%)
1	38 (3.8%)
2	8 (0.8%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___LOG	
0	907 (91%)
1	84 (8.4%)
2	7 (0.7%)

Query parameter or value	Descriptive statistics
3	2 (0.2%)
SELECT_n_of_non_aggr_func___DOW	
0	907 (91%)
1	80 (8.0%)
2	9 (0.9%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___YEAR	
0	875 (88%)
1	105 (11%)
2	17 (1.7%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func___ROUND	
0	939 (94%)
1	56 (5.6%)
2	5 (0.5%)
SELECT_n_of_non_aggr_func___MONTH	
0	862 (86%)
1	107 (11%)
2	18 (1.8%)
3	11 (1.1%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___ABS	
0	947 (95%)
1	46 (4.6%)
2	7 (0.7%)
SELECT_n_of_non_aggr_func___TRUNC	
0	956 (96%)
1	39 (3.9%)
2	3 (0.3%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.00, 2.00, 44.00] 3.11 / 3.98
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 9.00] 0.69 / 1.47
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 9.00] 0.73 / 1.50
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 10.00] 0.56 / 1.33
SELECT_n_of_aggr_func___SUM	
0	940 (94%)
1	45 (4.5%)
2	13 (1.3%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 10.00] 0.64 / 1.50
SELECT_n_of_aggr_func___AVG	
0	951 (95%)

Query parameter or value	Descriptive statistics
1	35 (3.5%)
2	12 (1.2%)
3	2 (0.2%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 11.00] 2.76 / 2.25
FROM_n_of_join_paths	
1	334 (33%)
2	331 (33%)
3	335 (34%)
FROM_n_of_super_joins___FULL	
0	779 (78%)
1	196 (20%)
2	25 (2.5%)
FROM_n_of_super_joins___RIGHT	
0	751 (75%)
1	229 (23%)
2	20 (2.0%)
FROM_n_of_super_joins___LEFT	
0	788 (79%)
1	188 (19%)
2	24 (2.4%)
FROM_n_of_joins___RIGHT	
0	240 (24%)
1	276 (28%)
2	192 (19%)
3	154 (15%)
4	88 (8.8%)
5	34 (3.4%)
6	12 (1.2%)
7	1 (0.1%)
8	3 (0.3%)
FROM_n_of_joins___INNER	
0	251 (25%)
1	271 (27%)
2	207 (21%)
3	145 (15%)
4	73 (7.3%)
5	34 (3.4%)
6	15 (1.5%)
7	4 (0.4%)
FROM_n_of_processed_rows	[5, 1,362,018,459, 4,590,055,357] 1,261,686,373 / 1,200,745,300
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.79 / 2.21
WHERE_n_of_attribs_of_type___character	
0	732 (73%)
1	203 (20%)
2	56 (5.6%)

Query parameter or value	Descriptive statistics
3	9 (0.9%)
WHERE_n_of_attribs_of_type__character_varying	
0	420 (42%)
1	340 (34%)
2	147 (15%)
3	71 (7.1%)
4	16 (1.6%)
5	5 (0.5%)
6	1 (0.1%)
WHERE_n_of_attribs_of_type__integer	
0	464 (46%)
1	324 (32%)
2	144 (14%)
3	48 (4.8%)
4	16 (1.6%)
5	4 (0.4%)
WHERE_n_of_attribs_of_type__numeric	
0	641 (64%)
1	256 (26%)
2	82 (8.2%)
3	20 (2.0%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	844 (84%)
1	132 (13%)
2	22 (2.2%)
3	2 (0.2%)
WHERE_n_of_pkey_attribs	
0	636 (64%)
1	280 (28%)
2	69 (6.9%)
3	15 (1.5%)
WHERE_n_of_connect_OR	
0	459 (46%)
1	174 (17%)
2	142 (14%)
3	91 (9.1%)
4	62 (6.2%)
5	46 (4.6%)
6	13 (1.3%)
7	12 (1.2%)
9	1 (0.1%)
WHERE_n_of_operators__greater_or_less	
0	401 (40%)
1	304 (30%)
2	194 (19%)

Query parameter or value	Descriptive statistics
3	59 (5.9%)
4	30 (3.0%)
5	10 (1.0%)
6	2 (0.2%)
WHERE_n_of_operators_in	
0	620 (62%)
1	277 (28%)
2	85 (8.5%)
3	17 (1.7%)
4	1 (0.1%)
WHERE_n_of_operators_like	
0	849 (85%)
1	137 (14%)
2	12 (1.2%)
3	2 (0.2%)
WHERE_n_of_operators_between	
0	600 (60%)
1	292 (29%)
2	89 (8.9%)
3	16 (1.6%)
4	3 (0.3%)
WHERE_n_of_non_aggr_func_FLOOR	
0	971 (97%)
1	28 (2.8%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_SQRT	
0	912 (91%)
1	81 (8.1%)
2	7 (0.7%)
WHERE_n_of_non_aggr_func_DAY	
0	974 (97%)
1	26 (2.6%)
WHERE_n_of_non_aggr_func_ABS	
0	966 (97%)
1	33 (3.3%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_LOG	
0	937 (94%)
1	59 (5.9%)
2	4 (0.4%)
WHERE_n_of_non_aggr_func_MONTH	
0	976 (98%)
1	23 (2.3%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_TRUNC	
0	967 (97%)

Query parameter or value	Descriptive statistics
1	33 (3.3%)
WHERE_n_of_non_aggr_func__YEAR	
0	978 (98%)
1	21 (2.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__ROUND	
0	975 (98%)
1	24 (2.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DOW	
0	980 (98%)
1	20 (2.0%)
WHERE_n_of_all_non_aggr_func	
0	711 (71%)
1	210 (21%)
2	68 (6.8%)
3	10 (1.0%)
4	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 10.00] 2.44 / 2.19
HAVING_n_of_main_predicates	
0	278 (28%)
1	89 (8.9%)
2	142 (14%)
3	173 (17%)
4	163 (16%)
5	94 (9.4%)
6	61 (6.1%)
HAVING_n_of_main_predicates__scalar_subquery	
0	418 (42%)
1	200 (20%)
2	176 (18%)
3	206 (21%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 19.0] 2.3 / 3.3
HAVING_n_of_processed_rows_by_subquery	[0, 62,000,135, 70,752,886,112] 2,998,864,248 / 6,901,963,991
ORDER_BY_n_of_columns	
0	310 (31%)
1	366 (37%)
2	238 (24%)
3	68 (6.8%)
4	14 (1.4%)
5	3 (0.3%)
6	1 (0.1%)
limit	[1, 482, 998] 496 / 278
offset	[0, 0, 996] 241 / 322