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

```
## Scenario: 1-3jp_rjjp__g_h0-3ss_h0-3nss

## Scale factor: 100 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)
```

Query parameter or value	Descriptive statistics	
SELECT_n_of_columns	[1, 6, 102] 10 / 14	
SELECT_n_of_non_aggr_funcRTRIM		
0	721~(72%)	
1	206 (21%)	
2	51 (5.1%)	
3	$14 \ (1.4\%)$	
4	6~(0.6%)	
5	1~(0.1%)	
6	1~(0.1%)	
SELECT_n_of_non_aggr_funcSUBSTR		
0	690~(69%)	
1	240 (24%)	
2	54 (5.4%)	
3	8 (0.8%)	
4	$3\;(0.3\%)$	

- for higher cardinality parameters: [min, median, max] mean / SD

Query parameter or value	Descriptive statistics
5	3 (0.3%)
6	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_funcLOG	•
0	900 (90%)
1	87 (8.7%)
2	9~(0.9%)
3	3~(0.3%)
4	1 (0.1%)
SELECT_n_of_non_aggr_funcLOWER	
0	699~(70%)
1	227~(23%)
2	55~(5.5%)
3	$16 \ (1.6\%)$
4	3~(0.3%)
SELECT_n_of_non_aggr_funcLTRIM	
0	718 (72%)
1	210 (21%)
2	52 (5.2%)
3	14 (1.4%)
4	3~(0.3%)
5	3~(0.3%)
SELECT_n_of_non_aggr_funcMONTH	
0	848~(85%)
1	117~(12%)
2	$21 \ (2.1\%)$
3	7 (0.7%)
4	3~(0.3%)
5	$3\;(0.3\%)$
6	1 (0.1%)
SELECT_n_of_non_aggr_funcSQRT	
0	850 (85%)
1	120 (12%)
2	$20 \ (2.0\%)$
3	4 (0.4%)
4	5~(0.5%)
5	1~(0.1%)
SELECT_n_of_non_aggr_funcUPPER	
0	694 (69%)
1	$223\ (22\%)$
2	61 (6.1%)
3	18 (1.8%)
4	2 (0.2%)
5	1 (0.1%)
6	1~(0.1%)
SELECT_n_of_non_aggr_funcFLOOR	

Query parameter or value	Descriptive statistics
0	945 (95%)
1	52 (5.2%)
2	1(0.1%)
3	1(0.1%)
4	1(0.1%)
SELECT_n_of_non_aggr_funcABS	, ,
0	944 (94%)
1	50 (5.0%)
2	5~(0.5%)
3	1 (0.1%)
SELECT_n_of_non_aggr_funcDOW	
0	880 (88%)
1	96 (9.6%)
2	10 (1.0%)
3	6 (0.6%)
4	6~(0.6%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_funcROUND	
0	944 (94%)
1	54 (5.4%)
2	2 (0.2%)
SELECT_n_of_non_aggr_funcYEAR	
0	865~(87%)
1	$106 \ (11\%)$
2	19~(1.9%)
3	4 (0.4%)
4	5 (0.5%)
5	1 (0.1%)
SELECT_n_of_non_aggr_funcTRUNC	(220)
0	945 (95%)
1	52 (5.2%)
2	3~(0.3%)
SELECT_n_of_non_aggr_funcDAY	224 (2274)
0	921 (92%)
1	67 (6.7%)
2	10 (1.0%)
4	2 (0.2%)
SELECT_n_of_all_non_aggr_func	[0.0, 2.0, 38.0] 3.2 / 4.5
SELECT_n_of_aggr_funcAVG	000 (0004)
0	926 (93%)
1	52 (5.2%)
2	18 (1.8%)
3	$4 \; (0.4\%)$
SELECT_n_of_aggr_funcMAX	740 (7407)
0	740 (74%)

Query parameter or value	Descriptive statistics
1	81 (8.1%)
2	71 (7.1%)
3	34 (3.4%)
4	31 (3.1%)
5	$22\ (2.2\%)$
6	14~(1.4%)
7	6~(0.6%)
8	1 (0.1%)
SELECT_n_of_aggr_funcMIN	$[0.00, 0.00, 11.00] \ 0.61 \ / \ 1.42$
SELECT_n_of_aggr_funcCOUNT	[0.00, 0.00, 10.00] $[0.59]$ / 1.41
SELECT_n_of_aggr_funcCOUNT_DISTINCT	[0.00, 0.00, 11.00] $[0.65]$ $[0.44]$
SELECT_n_of_aggr_funcSUM	
0	940 (94%)
1	47 (4.7%)
2	13~(1.3%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 11.00] 2.71 / 2.27
FROM_n_of_join_paths	
1	345~(35%)
2	327 (33%)
3	328 (33%)
FROM_n_of_super_joinsRIGHT	
0	770 (77%)
1	$212\ (21\%)$
2	18 (1.8%)
FROM_n_of_super_joinsLEFT	
0	803 (80%)
1	178 (18%)
2	19~(1.9%)
FROM_n_of_super_joinsFULL	
0	781 (78%)
1	195~(20%)
2	24 (2.4%)
FROM_n_of_joinsINNER	
0	272~(27%)
1	250~(25%)
2	222~(22%)
3	121~(12%)
4	73 (7.3%)
5	42 (4.2%)
6	10 (1.0%)
7	6~(0.6%)
8	4~(0.4%)
$FROM_n_of_joins__RIGHT$	
0	243~(24%)
1	269(27%)

138 (14%) 4	Query parameter or value	Descriptive statistics
4 75 (7.5%) 5 25 (2.5%) 6 16 (1.6%) 7 8 (0.8%) 8 2 (0.2%) FROM_n of processed_rows [5, 681,037,927, 2,295,113,786] 647,545,503 / 600,842,516 WHERE_n_of_predicates [0.00, 3.00, 10.00] 2.82 / 2.20 WHERE_n_of_attribs_of_type_character_varying 430 (43%) 1 325 (33%) 2 154 (15%) 3 70 (7.0%) 4 18 (1.8%) 5 3 (0.3%) WHERE_n_of_attribs_of_type_integer 443 (44%) 0 443 (44%) 1 304 (30%) 2 185 (19%) 3 48 (4.8%) 4 44 (44%) 1 3 (0.3%) WHERE_n_of_attribs_of_type_numeric 649 (65%) 0 649 (65%) 1 257 (26%) 2 79 (7.9%) 3 10.1%) WHERE_n_of_attribs_of_type_date 846 (85%) 1 126 (13%) 2 26 (2.6%) 3 20.2%) WHERE	3	138 (14%)
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	4	75 (7.5%)
7 8 (0.8%) 8 2 (0.2%) FROM_n_of_processed_rows [5, 681,037,927, 2,295,113,786] 647,545,503 / 600,842,516 WHERE_n_of_predicates [0.00, 3.00, 10.00] 2.82 / 2.20 WHERE_n_of_attribs_of_type_character varying 0 430 (43%) 1 325 (33%) 2 154 (15%) 3 70 (7.0%) 4 18 (1.8%) 5 3 (0.3%) WHERE_n_of_attribs_of_type_integer 443 (44%) 0 43 (30%) 2 185 (19%) 3 48 (4.8%) 4 17 (1.7%) 5 3 (0.3%) WHERE_n_of_attribs_of_type_numeric 649 (65%) 0 649 (65%) 1 1 (0.1%) WHERE_n_of_attribs_of_type_date 846 (85%) 1 1 (0.1%) WHERE_n_of_attribs_of_type_character 724 (72%) 0 724 (72%) 1 22 (0.2%) WHERE_n_of_attribs_of_type_character 724 (72%) 1 222 (22%) 2 43 (4.3%) 3<		$25\ (2.5\%)$
8 C (0.2%) FROM_n_of_processed_rows [5, 681,037,927, 2,295,113,786] 647,545,503 / 600,842,516 WHERE_n_of_predicates [0.00, 3.00, 10.00] 2.82 / 2.20 WHERE_n_of_attribs_of_typecharacter_varying 430 (43%) 0 430 (43%) 1 325 (33%) 2 154 (15%) 3 70 (7.0%) 4 18 (1.8%) 5 3 (0.3%) WHERE_n_of_attribs_of_typeinteger 443 (44%) 1 304 (30%) 2 185 (19%) 3 48 (4.8%) 4 17 (1.7%) 5 3 (0.33%) WHERE_n_of_attribs_of_typenumeric 649 (65%) 1 257 (26%) 2 79 (7.9%) 3 14 (1.4%) 4 1 (0.1%) WHERE_n_of_attribs_of_typedate 846 (85%) 1 126 (13%) 2 26 (2.6%) 3 2 (0.2%) WHERE_n_of_attribs_of_typecharacter 724 (72%) 1 222 (22%) <	6	16(1.6%)
FROM_n_of_processed_rows	7	8 (0.8%)
WHERE_n_of_predicates WHERE_n_of_attribs_of_typecharacter_varying 0	8	2(0.2%)
WHERE_n_of_attribs_of_typecharacter_varying 0	$FROM_n_of_processed_rows$	
WHERE_n_of_attribs_of_typecharacter_varying 0	WHERE n of predicates	, ,
0 430 (43%) 1 325 (33%) 2 154 (15%) 3 70 (7.0%) 4 18 (1.8%) 5 3 (0.3%) WHERE_n_of_attribs_of_type_integer 443 (44%) 1 304 (30%) 2 185 (19%) 3 48 (4.8%) 4 17 (1.7%) 5 3 (0.3%) WHERE_n_of_attribs_of_type_numeric 649 (65%) 0 649 (65%) 1 257 (26%) 2 79 (7.9%) 3 14 (1.4%) 4 1 (0.1%) WHERE_n_of_attribs_of_type_date 846 (85%) 0 846 (85%) 1 126 (13%) 2 26 (2.6%) 3 2 (0.2%) WHERE_n_of_attribs_of_type_character 724 (72%) 1 222 (22%) 2 43 (4.3%) 3 10 (1.0%) 4 1 (0.1%) WHERE_n_of_pkey_attribs 606 (61%)		
1 325 (33%) 2 154 (15%) 3 70 (7.0%) 4 18 (1.8%) 5 3 (0.3%) WHERE_n_of_attribs_of_typeinteger WHERE_n_of_attribs_of_typeinteger 0 443 (44%) 1 304 (30%) 2 185 (19%) 3 48 (4.8%) 4 17 (1.7%) 5 3 (0.3%) WHERE_n_of_attribs_of_typenumeric 649 (65%) 1 257 (26%) 2 79 (7.9%) 3 14 (1.4%) 4 1 (0.1%) WHERE_n_of_attribs_of_typedate 846 (85%) 1 126 (13%) 2 26 (2.6%) 3 2 (0.2%) WHERE_n_of_attribs_of_typecharacter 724 (72%) 1 222 (22%) 2 43 (4.3%) 3 10 (1.0%) WHERE_n_of_pkey_attribs 606 (61%)		
$\begin{array}{cccccccccccccccccccccccccccccccccccc$, ,
$\begin{array}{cccccccccccccccccccccccccccccccccccc$, ,
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$, ,
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	4	, ,
$\begin{array}{cccccccccccccccccccccccccccccccccccc$, ,
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	WHERE n of attribs of type integer	` '
$\begin{array}{cccccccccccccccccccccccccccccccccccc$		
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	1	,
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	2	, ,
$\begin{array}{cccccccccccccccccccccccccccccccccccc$, ,
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	4	· ,
$\begin{array}{cccccccccccccccccccccccccccccccccccc$		· ,
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	WHERE n of attribs of type numeric	, ,
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	0	649~(65%)
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	1	$257\ (26\%)$
4 1 (0.1%) WHERE_n_of_attribs_of_typedate 846 (85%) 0 846 (13%) 2 26 (2.6%) 3 2 (0.2%) WHERE_n_of_attribs_of_typecharacter 724 (72%) 1 222 (22%) 2 43 (4.3%) 3 10 (1.0%) 4 1 (0.1%) WHERE_n_of_pkey_attribs 606 (61%)	2	79 (7.9%)
WHERE_n_of_attribs_of_typedate 0	3	14 (1.4%)
0 846 (85%) 1 126 (13%) 2 26 (2.6%) 3 2 (0.2%) WHERE_n_of_attribs_of_typecharacter 724 (72%) 1 222 (22%) 2 43 (4.3%) 3 10 (1.0%) 4 1 (0.1%) WHERE_n_of_pkey_attribs 606 (61%)	4	1 (0.1%)
1 126 (13%) 2 26 (2.6%) 3 2 (0.2%) WHERE_n_of_attribs_of_typecharacter 0 724 (72%) 1 222 (22%) 2 43 (4.3%) 3 10 (1.0%) 4 1 (0.1%) WHERE_n_of_pkey_attribs 606 (61%)	WHERE_n_of_attribs_of_typedate	
26 (2.6%) 3	0	846 (85%)
3 2 (0.2%) WHERE_n_of_attribs_of_typecharacter 724 (72%) 1 222 (22%) 2 43 (4.3%) 3 10 (1.0%) 4 1 (0.1%) WHERE_n_of_pkey_attribs 606 (61%)	1	126 (13%)
WHERE_n_of_attribs_of_typecharacter 0	2	$26\ (2.6\%)$
0 724 (72%) 1 222 (22%) 2 43 (4.3%) 3 10 (1.0%) 4 1 (0.1%) WHERE_n_of_pkey_attribs 606 (61%) 0 606 (61%)	3	2~(0.2%)
1 222 (22%) 2 43 (4.3%) 3 10 (1.0%) 4 1 (0.1%) WHERE_n_of_pkey_attribs 0 606 (61%)	WHERE_n_of_attribs_of_typecharact	ter
2 43 (4.3%) 3 10 (1.0%) 4 1 (0.1%) WHERE_n_of_pkey_attribs 0 606 (61%)	0	724~(72%)
3 10 (1.0%) 4 1 (0.1%) WHERE_n_of_pkey_attribs 0 606 (61%)	1	222~(22%)
4 1 (0.1%) WHERE_n_of_pkey_attribs 0 606 (61%)	2	, ,
WHERE_n_of_pkey_attribs 0 606 (61%)	3	$10 \ (1.0\%)$
0 606 (61%)	4	1 (0.1%)
` '		
1 282 (28%)		606 (61%)
202 (2070)	1	282 (28%)
2 101 (10%)	2	101 (10%)
9 (0.9%)	3	$9\;(0.9\%)$
(0.2%)	4	2~(0.2%)

Query parameter or value	Descriptive statistics
WHERE_n_of_connect_OR	
θ	455~(46%)
1	167 (17%)
2	$140 \ (14\%)$
3	114 (11%)
4	65~(6.5%)
5	$30 \ (3.0\%)$
6	$23\ (2.3\%)$
7	5~(0.5%)
8	1 (0.1%)
WHERE_n_of_operatorsgreater_or_less	,
0	418 (42%)
1	298 (30%)
2	162 (16%)
3	74 (7.4%)
4	$34\ (3.4\%)$
<i>.</i> 5	11 (1.1%)
6	2 (0.2%)
$\overline{\gamma}$	1 (0.1%)
WHERE n of operators between	= (8.=70)
0	609 (61%)
1	278 (28%)
2	94 (9.4%)
~ 3	17 (1.7%)
4	2 (0.2%)
WHERE_n_of_operatorsin	= (0.270)
0	603~(60%)
1	294 (29%)
2	81 (8.1%)
~ 3	19 (1.9%)
4	3 (0.3%)
WHERE_n_of_operatorslike	0 (0.070)
0	819 (82%)
1	163 (16%)
2	17 (1.7%)
~ 3	1 (0.1%)
WHERE_n_of_non_aggr_funcFLOOR	1 (0.170)
0	963 (96%)
1	36 (3.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_funcLOG	1 (0.1/0)
0	946 (95%)
1	54 (5.4%)
WHERE_n_of_non_aggr_funcSQRT	9 4 (9.4/0)
	012 (01%)
0	913 (91%)
1	80 (8.0%)

Query parameter or value	Descriptive statistics
2	7 (0.7%)
WHERE_n_of_non_aggr_funcYEAR	,
0	976 (98%)
1	24(2.4%)
WHERE_n_of_non_aggr_funcDOW	,
0	977 (98%)
1	$22\ (2.2\%)$
2	1 (0.1%)
WHERE_n_of_non_aggr_funcABS	
0	974 (97%)
1	25~(2.5%)
2	1 (0.1%)
WHERE_n_of_non_aggr_funcROUND	` ,
0	974 (97%)
1	$25 \ (2.5\%)$
2	1 (0.1%)
WHERE_n_of_non_aggr_funcMONTH	,
0	970 (97%)
1	$29 \ (2.9\%)$
2	1 (0.1%)
WHERE_n_of_non_aggr_funcDAY	(- , ,)
0	980 (98%)
1	$20\ (2.0\%)$
WHERE_n_of_non_aggr_funcTRUNC	
0 = 0	977 (98%)
1	$22\ (2.2\%)$
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	,
0	715~(72%)
1	$222\ (22\%)$
2	51 (5.1%)
3	9 (0.9%)
4	3~(0.3%)
GROUP_BY_n_of_columns	` ,
0	262~(26%)
1	186 (19%)
2	151 (15%)
3	135 (14%)
4	108 (11%)
5	79 (7.9%)
6	47 (4.7%)
7	$22\ (2.2\%)$
8	10 (1.0%)
HAVING_n_of_main_predicates	[0.00, 3.00, 9.00] $[0.00, 3.00, 9.00]$ $[0.00, 3.00, 9.00]$
-	L / / J /
HAVING_n_of_main_predicatesscalar_subquery	

Query parameter or value	Descriptive statistics
1	171 (17%)
2	175 (18%)
3	177 (18%)
HAVING_n_of_main_predicatesnon_scalar_s	subquery
0	469~(47%)
1	171 (17%)
2	162~(16%)
3	198 (20%)
HAVING_n_of_subqueriesscalar_subquery	$[0.0, 1.0, 24.0] \ 2.3 \ / \ 3.5$
HAVING_n_of_subqueriesnon_scalar_subque	ery
0	469 (47%)
1	171 (17%)
2	162 (16%)
3	198 (20%)
HAVING_n_of_processed_rows_by_subqueries;	361,000,150, 33,777,819,296] 2,145,162,434 /
	3,979,348,136
ORDER_BY_n_of_columns	
0	306 (31%)
1	$368 \; (37\%)$
2	247~(25%)
3	54 (5.4%)
4	12~(1.2%)
5	7~(0.7%)
6	4 (0.4%)
7	2~(0.2%)
limit	[2, 481, 1,000] 490 / 291
offset	[0, 14, 1,000] 254 / 328