

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

Scenario: 1-1jp__g__h0ss_h0nss

Scale factor: 1000 GB

The 1000-query pack was generated on: 2023-11-04

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	[1.0, 4.0, 42.0] 5.8 / 5.5
SELECT_n_of_non_aggr_func__ABS	
0	960 (96%)
1	38 (3.8%)
2	2 (0.2%)
SELECT_n_of_non_aggr_func__LOWER	
0	776 (78%)
1	195 (20%)
2	26 (2.6%)
3	2 (0.2%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__LTRIM	
0	790 (79%)
1	183 (18%)
2	25 (2.5%)

Query parameter or value	Descriptive statistics
3	2 (0.2%)
SELECT__n_of__non__aggr__func__RTRIM	
0	787 (79%)
1	188 (19%)
2	22 (2.2%)
3	2 (0.2%)
4	1 (0.1%)
SELECT__n_of__non__aggr__func__SUBSTR	
0	771 (77%)
1	194 (19%)
2	30 (3.0%)
3	5 (0.5%)
SELECT__n_of__non__aggr__func__SQRT	
0	901 (90%)
1	92 (9.2%)
2	7 (0.7%)
SELECT__n_of__non__aggr__func__LOG	
0	933 (93%)
1	60 (6.0%)
2	6 (0.6%)
3	1 (0.1%)
SELECT__n_of__non__aggr__func__DOW	
0	948 (95%)
1	46 (4.6%)
2	4 (0.4%)
3	2 (0.2%)
SELECT__n_of__non__aggr__func__MONTH	
0	913 (91%)
1	69 (6.9%)
2	16 (1.6%)
3	2 (0.2%)
SELECT__n_of__non__aggr__func__UPPER	
0	788 (79%)
1	186 (19%)
2	23 (2.3%)
3	3 (0.3%)
SELECT__n_of__non__aggr__func__DAY	
0	949 (95%)
1	45 (4.5%)
2	5 (0.5%)
3	1 (0.1%)
SELECT__n_of__non__aggr__func__FLOOR	
0	967 (97%)
1	31 (3.1%)
2	2 (0.2%)
SELECT__n_of__non__aggr__func__YEAR	

Query parameter or value	Descriptive statistics
0	928 (93%)
1	61 (6.1%)
2	9 (0.9%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func__ROUND	
0	975 (98%)
1	25 (2.5%)
SELECT_n_of_non_aggr_func__TRUNC	
0	969 (97%)
1	30 (3.0%)
2	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.00, 1.00, 18.00] 1.87 / 2.15
SELECT_n_of_aggr_func__MIN	
0	780 (78%)
1	116 (12%)
2	62 (6.2%)
3	26 (2.6%)
4	10 (1.0%)
5	3 (0.3%)
6	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_aggr_func__AVG	
0	955 (96%)
1	34 (3.4%)
2	11 (1.1%)
SELECT_n_of_aggr_func__MAX	
0	768 (77%)
1	94 (9.4%)
2	70 (7.0%)
3	40 (4.0%)
4	15 (1.5%)
5	7 (0.7%)
6	4 (0.4%)
7	2 (0.2%)
SELECT_n_of_aggr_func__COUNT_DISTINCT	
0	804 (80%)
1	93 (9.3%)
2	53 (5.3%)
3	28 (2.8%)
4	12 (1.2%)
5	8 (0.8%)
6	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_aggr_func__SUM	
0	958 (96%)
1	34 (3.4%)

Query parameter or value	Descriptive statistics
2	8 (0.8%)
SELECT_n_of_aggr_func__COUNT	
0	759 (76%)
1	115 (12%)
2	70 (7.0%)
3	32 (3.2%)
4	15 (1.5%)
5	7 (0.7%)
6	2 (0.2%)
SELECT_n_of_all_aggr_func	
0	208 (21%)
1	276 (28%)
2	227 (23%)
3	151 (15%)
4	74 (7.4%)
5	43 (4.3%)
6	16 (1.6%)
7	3 (0.3%)
8	2 (0.2%)
FROM_n_of_join_paths	
1	1,000 (100%)
FROM_n_of_joins__INNER	
0	504 (50%)
1	320 (32%)
2	149 (15%)
3	23 (2.3%)
4	4 (0.4%)
FROM_n_of_joins__RIGHT	
0	481 (48%)
1	341 (34%)
2	138 (14%)
3	33 (3.3%)
4	7 (0.7%)
FROM_n_of_processed_rows	[5, 1,000,000,000, 7,649,989,739] 2,914,486,288 / 3,165,754,122
WHERE_n_of_predicates	
0	247 (25%)
1	234 (23%)
2	212 (21%)
3	156 (16%)
4	90 (9.0%)
5	50 (5.0%)
6	11 (1.1%)
WHERE_n_of_attrihs_of_type__character	
0	834 (83%)
1	146 (15%)

Query parameter or value	Descriptive statistics
<i>2</i>	19 (1.9%)
<i>3</i>	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
<i>0</i>	893 (89%)
<i>1</i>	91 (9.1%)
<i>2</i>	16 (1.6%)
WHERE_n_of_attribs_of_type__integer	
<i>0</i>	596 (60%)
<i>1</i>	305 (31%)
<i>2</i>	85 (8.5%)
<i>3</i>	13 (1.3%)
<i>4</i>	1 (0.1%)
WHERE_n_of_attribs_of_type__numeric	
<i>0</i>	723 (72%)
<i>1</i>	227 (23%)
<i>2</i>	38 (3.8%)
<i>3</i>	11 (1.1%)
<i>4</i>	1 (0.1%)
WHERE_n_of_attribs_of_type__character_varying	
<i>0</i>	527 (53%)
<i>1</i>	340 (34%)
<i>2</i>	108 (11%)
<i>3</i>	23 (2.3%)
<i>4</i>	1 (0.1%)
<i>5</i>	1 (0.1%)
WHERE_n_of_pkey_attribs	
<i>0</i>	732 (73%)
<i>1</i>	226 (23%)
<i>2</i>	40 (4.0%)
<i>3</i>	2 (0.2%)
WHERE_n_of_connect_OR	
<i>0</i>	612 (61%)
<i>1</i>	213 (21%)
<i>2</i>	94 (9.4%)
<i>3</i>	63 (6.3%)
<i>4</i>	15 (1.5%)
<i>5</i>	3 (0.3%)
WHERE_n_of_operators__between	
<i>0</i>	737 (74%)
<i>1</i>	222 (22%)
<i>2</i>	37 (3.7%)
<i>3</i>	4 (0.4%)
WHERE_n_of_operators__greater_or_less	
<i>0</i>	538 (54%)
<i>1</i>	305 (31%)
<i>2</i>	117 (12%)

Query parameter or value	Descriptive statistics
3	32 (3.2%)
4	7 (0.7%)
5	1 (0.1%)
WHERE_n_of_operators_in	
0	731 (73%)
1	220 (22%)
2	44 (4.4%)
3	5 (0.5%)
WHERE_n_of_operators_like	
0	870 (87%)
1	123 (12%)
2	7 (0.7%)
WHERE_n_of_non_aggr_func_FLOOR	
0	970 (97%)
1	29 (2.9%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_MONTH	
0	982 (98%)
1	18 (1.8%)
WHERE_n_of_non_aggr_func_SQRT	
0	950 (95%)
1	50 (5.0%)
WHERE_n_of_non_aggr_func_DAY	
0	976 (98%)
1	24 (2.4%)
WHERE_n_of_non_aggr_func_YEAR	
0	982 (98%)
1	17 (1.7%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_DOW	
0	987 (99%)
1	13 (1.3%)
WHERE_n_of_non_aggr_func_ABS	
0	977 (98%)
1	23 (2.3%)
WHERE_n_of_non_aggr_func_ROUND	
0	984 (98%)
1	15 (1.5%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_TRUNC	
0	977 (98%)
1	23 (2.3%)
WHERE_n_of_non_aggr_func_LOG	
0	973 (97%)
1	27 (2.7%)
WHERE_n_of_all_non_aggr_func	

Query parameter or value	Descriptive statistics
<i>0</i>	798 (80%)
<i>1</i>	165 (17%)
<i>2</i>	32 (3.2%)
<i>3</i>	4 (0.4%)
<i>4</i>	1 (0.1%)
GROUP_BY_n_of_columns	
<i>0</i>	277 (28%)
<i>1</i>	272 (27%)
<i>2</i>	215 (22%)
<i>3</i>	121 (12%)
<i>4</i>	75 (7.5%)
<i>5</i>	31 (3.1%)
<i>6</i>	9 (0.9%)
HAVING_n_of_main_predicates	
<i>0</i>	451 (45%)
<i>1</i>	193 (19%)
<i>2</i>	167 (17%)
<i>3</i>	189 (19%)
ORDER_BY_n_of_columns	
<i>0</i>	301 (30%)
<i>1</i>	468 (47%)
<i>2</i>	197 (20%)
<i>3</i>	27 (2.7%)
<i>4</i>	4 (0.4%)
<i>5</i>	3 (0.3%)
limit	[1, 498, 999] 499 / 297
offset	[0, 5, 997] 256 / 328