

Real				
	native.country	marital.status	SUM(fnlwgt)	COUNT(*)
197	Thailand	Divorced	876803	3
129	Japan	Divorced	2928035	13
95	Hungary	Divorced	332719	2
196	Taiwan	Never-married	5091872	25
70	Greece	Divorced	563784	3

Synthetic				
	native.country	marital.status	SUM(fnlwgt)	COUNT(*)
197	Thailand	Divorced	904986.196604	5.000000
129	Japan	Divorced	44034055.207424	247.000000
95	Hungary	Divorced	29644820.469457	163.000000
196	Taiwan	Never-married	6998547.522023	37.000000
70	Greece	Divorced	nan	nan

SQL for Real:
SELECT `native.country`,`marital.status`,SUM(fnlwgt), COUNT(*) FROM C1 GROUP BY `native.country`,`marital.status`
Resulted in 221 records

SQL for Synthetic:
SELECT `native.country`,`marital.status`,SUM(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY `native.country`,`marital.status`
Resulted in 196 records

Normalized Euclidean distance for (fnlwgt): 13.19
Hellinger Distance: 0.4

=====

Real					
	relationship	race	occupation	MAX(age)	COUNT(*)
227	Own-child	Black	Machine-op-inspct	52	60
149	Other-relative	Asian-Pac-Islander	Handlers-cleaners	53	4
80	Not-in-family	Asian-Pac-Islander	Adm-clerical	71	44
49	Husband	Other	Tech-support	40	2
276	Unmarried	Asian-Pac-Islander	?	72	8

Synthetic					
	relationship	race	occupation	MAX(age)	COUNT(*)
227	Own-child	Black	Machine-op-inspct	57.000000	55.000000
149	Other-relative	Asian-Pac-Islander	Handlers-cleaners	28.000000	2.000000
80	Not-in-family	Asian-Pac-Islander	Adm-clerical	70.000000	32.000000
49	Husband	Other	Tech-support	27.000000	1.000000
276	Unmarried	Asian-Pac-Islander	?	27.000000	1.000000

SQL for Real:
SELECT relationship,race,occupation,MAX(age), COUNT(*) FROM C1 GROUP BY relationship,race,occupation
Resulted in 387 records

SQL for Synthetic:
SELECT relationship,race,occupation,MAX(age), COUNT(*) FROM C1_syn_06 GROUP BY relationship,race,occupation
Resulted in 376 records

Normalized Euclidean distance for (age): 19.0
Hellinger Distance: 0.064

=====

Real											
	education	native.country	occupation	race	relationship	sex	workclass	income	marital.status	MIN(`hours.per.week`)	COUNT(*)
4653	Assoc-voc	Puerto-Rico	Prof-specialty	White	Husband	Male	Private	<=50K	Married-civ-spouse	nan	nan
2495	7th-8th	Puerto-Rico	Machine-op-inspct	Other	Unmarried	Female	Private	<=50K	Widowed	nan	nan
944	11th	Puerto-Rico	?	White	Not-in-family	Male	?	<=50K	Divorced	nan	nan
4102	Assoc-acdm	United-States	Sales	Other	Own-child	Female	Private	<=50K	Never-married	nan	nan
15752	Some-college	Puerto-Rico	Tech-support	White	Not-in-family	Female	State-gov	<=50K	Never-married	nan	nan

Synthetic											
	education	native.country	occupation	race	relationship	sex	workclass	income	marital.status	MIN(`hours.per.week`)	COUNT(*)
4653	Assoc-voc	Puerto-Rico	Prof-specialty	White	Husband	Male	Private	<=50K	Married-civ-spouse	40.189674	7
2495	7th-8th	Puerto-Rico	Machine-op-inspct	Other	Unmarried	Female	Private	<=50K	Widowed	40.002161	1
944	11th	Puerto-Rico	?	White	Not-in-family	Male	?	<=50K	Divorced	39.993443	2
4102	Assoc-acdm	United-States	Sales	Other	Own-child	Female	Private	<=50K	Never-married	39.552374	1
15752	Some-college	Puerto-Rico	Tech-support	White	Not-in-family	Female	State-gov	<=50K	Never-married	39.957147	2

SQL for Real:
SELECT education,`native.country`,occupation,race,relationship,sex,workclass,income,`marital.status`,MIN(`hours.per.week`), COUNT(*) FROM C1 GROUP BY education,`native.country`,occupation,race,relationship,sex,workclass,income,`marital.status`
Resulted in 12664 records

SQL for Synthetic:
SELECT education,`native.country`,occupation,race,relationship,sex,workclass,income,`marital.status`,MIN(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY education,`native.country`,occupation,race,relationship,sex,workclass,income,`marital.status`
Resulted in 17127 records

Normalized Euclidean distance for (`hours.per.week`): 64.61

Hellinger Distance: 0.252

=====

Real										
	workclass	race	occupation	income	education	marital.status	native.country	sex	MAX(capital)	COUNT(*)
8734	Private	White	Priv-house-serv	<=50K	Bachelors	Never-married	United-States	Female	0.000000	2.000000
8077	Private	White	Other-service	<=50K	10th	Divorced	Cuba	Female	nan	nan
2000	Local-gov	White	Adm-clerical	<=50K	9th	Divorced	United-States	Female	nan	nan
7853	Private	White	Machine-op-inspct	<=50K	HS-grad	Divorced	Ireland	Male	nan	nan
5016	Private	Other	Adm-clerical	<=50K	HS-grad	Never-married	Thailand	Female	nan	nan

Synthetic										
	workclass	race	occupation	income	education	marital.status	native.country	sex	MAX(capital)	COUNT(*)
8734	Private	White	Priv-house-serv	<=50K	Bachelors	Never-married	United-States	Female	75.016811	1
8077	Private	White	Other-service	<=50K	10th	Divorced	Cuba	Female	38.575525	1
2000	Local-gov	White	Adm-clerical	<=50K	9th	Divorced	United-States	Female	31.430604	1
7853	Private	White	Machine-op-inspct	<=50K	HS-grad	Divorced	Ireland	Male	0.638358	1
5016	Private	Other	Adm-clerical	<=50K	HS-grad	Never-married	Thailand	Female	-0.306732	1

SQL for Real:
SELECT workclass,race,occupation,income,education,`marital.status`,`native.country`,sex,MAX(capital), COUNT(*) FROM C1 GROUP BY workclass,race,occupation,income,education,`marital.status`,`native.country`,sex

Resulted in 10025 records

SQL for Synthetic:
SELECT workclass,race,occupation,income,education,`marital.status`,`native.country`,sex,MAX(capital), COUNT(*) FROM C1_syn_06 GROUP BY workclass,race,occupation,income,education,`marital.status`,`native.country`,sex

Resulted in 13875 records

Normalized Euclidean distance for (capital): 59.76

Hellinger Distance: 0.252

=====

Real				
	native.country	workclass	MIN(age)	COUNT(*)
135	Jamaica	Private	18	87
59	England	Self-emp-not-inc	18	13
133	Jamaica	Federal-gov	48	1
113	India	State-gov	24	17
151	Mexico	Federal-gov	32	4

Synthetic				
	native.country	workclass	MIN(age)	COUNT(*)
135	Jamaica	Private	31.000000	42.000000
59	England	Self-emp-not-inc	27.000000	1.000000
133	Jamaica	Federal-gov	57.000000	2.000000
113	India	State-gov	28.000000	2.000000
151	Mexico	Federal-gov	38.000000	1.000000

SQL for Real:
SELECT `native.country`,workclass,MIN(age), COUNT(*) FROM C1 GROUP BY `native.country`,workclass

Resulted in 248 records

SQL for Synthetic:
SELECT `native.country`,workclass,MIN(age), COUNT(*) FROM C1_syn_06 GROUP BY `native.country`,workclass

Resulted in 207 records

Normalized Euclidean distance for (age): 13.56

Hellinger Distance: 0.394

=====

Real								
	occupation	relationship	race	education	workclass	sex	AVG(fnlwgt)	COUNT(*)
1324	Craft-repair	Unmarried	Asian-Pac-Islander	Masters	Private	Male	180239.000000	1
387	Adm-clerical	Not-in-family	Amer-Indian-Eskimo	Bachelors	Private	Male	190319.000000	1
2529	Handlers-cleaners	Own-child	White	Bachelors	Private	Male	198285.812500	16
1962	Exec-managerial	Wife	Asian-Pac-Islander	Bachelors	Self-emp-not-inc	Female	163559.000000	1
5240	Tech-support	Husband	White	HS-grad	Private	Male	185233.546875	64

Synthetic								
	occupation	relationship	race	education	workclass	sex	AVG(fnlwgt)	COUNT(*)
1324	Craft-repair	Unmarried	Asian-Pac-Islander	Masters	Private	Male	nan	nan
387	Adm-clerical	Not-in-family	Amer-Indian-Eskimo	Bachelors	Private	Male	179622.253562	1.000000
2529	Handlers-cleaners	Own-child	White	Bachelors	Private	Male	186270.115320	16.000000
1962	Exec-managerial	Wife	Asian-Pac-Islander	Bachelors	Self-emp-not-inc	Female	nan	nan
5240	Tech-support	Husband	White	HS-grad	Private	Male	178073.756101	72.000000

SQL for Real:
SELECT occupation,relationship,race,education,workclass,sex,AVG(fnlwgt), COUNT(*) FROM C1 GROUP BY occupation,relationship,race,education,workclass,sex

Resulted in 5820 records

SQL for Synthetic:

SELECT occupation,relationship,race,education,workclass,sex,AVG(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY occupation,relationship,race,education,workclass,sex

Resulted in 5511 records

Normalized Euclidean distance for (fnlwgt): 59.74

Hellinger Distance: 0.125

=====

Real				
	workclass	native.country	AVG(`hours.per.week`)	COUNT(*)
22	?	Japan	38.333333	3
229	State-gov	India	40.058824	17
216	Self-emp-not-inc	Yugoslavia	44.500000	4
14	?	Honduras	17.000000	1
70	Local-gov	China	38.333333	3

Synthetic				
	workclass	native.country	AVG(`hours.per.week`)	COUNT(*)
22	?	Japan	39.972852	55.000000
229	State-gov	India	40.015170	2.000000
216	Self-emp-not-inc	Yugoslavia	40.012129	4.000000
14	?	Honduras	nan	nan
70	Local-gov	China	39.993824	5.000000

SQL for Real:
SELECT workclass,`native.country`,AVG(`hours.per.week`), COUNT(*) FROM C1 GROUP BY workclass,`native.country`

Resulted in 248 records

SQL for Synthetic:
SELECT workclass,`native.country`,AVG(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY workclass,`native.country`

Resulted in 207 records

Normalized Euclidean distance for (`hours.per.week`): 13.56

Hellinger Distance: 0.394

=====

Real											
	sex	relationship	native.country	workclass	marital.status	occupation	race	income	education	MIN(capital)	COUNT(*)
9893	Male	Husband	Puerto-Rico	Private	Married-civ-spouse	Transport-moving	White	<=50K	Assoc-acdm	nan	nan
15884	Male	Own-child	United-States	Private	Married-civ-spouse	Sales	White	<=50K	Bachelors	nan	nan
15011	Male	Own-child	Columbia	Self-emp-not-inc	Never-married	Transport-moving	White	<=50K	HS-grad	nan	nan
3820	Female	Own-child	Taiwan	Private	Married-civ-spouse	Prof-specialty	White	<=50K	Assoc-voc	nan	nan
8910	Male	Husband	Japan	Self-emp-not-inc	Married-civ-spouse	Prof-specialty	White	<=50K	Prof-school	nan	nan

Synthetic											
	sex	relationship	native.country	workclass	marital.status	occupation	race	income	education	MIN(capital)	COUNT(*)
9893	Male	Husband	Puerto-Rico	Private	Married-civ-spouse	Transport-moving	White	<=50K	Assoc-acdm	-65.203050	4
15884	Male	Own-child	United-States	Private	Married-civ-spouse	Sales	White	<=50K	Bachelors	42.092501	1
15011	Male	Own-child	Columbia	Self-emp-not-inc	Never-married	Transport-moving	White	<=50K	HS-grad	-14.215888	1
3820	Female	Own-child	Taiwan	Private	Married-civ-spouse	Prof-specialty	White	<=50K	Assoc-voc	65.317075	1
8910	Male	Husband	Japan	Self-emp-not-inc	Married-civ-spouse	Prof-specialty	White	<=50K	Prof-school	-37.773696	4

SQL for Real:
SELECT sex,relationship,`native.country`,workclass,`marital.status`,occupation,race,income,education,MIN(capital), COUNT(*) FROM C1 GROUP BY sex,relationship,`native.country`,workclass,`marital.status`,occupation,race,income,education

Resulted in 12664 records

SQL for Synthetic:
SELECT sex,relationship,`native.country`,workclass,`marital.status`,occupation,race,income,education,MIN(capital), COUNT(*) FROM C1_syn_06 GROUP BY sex,relationship,`native.country`,workclass,`marital.status`,occupation,race,income,education

Resulted in 17127 records

Normalized Euclidean distance for (capital): 64.61

Hellinger Distance: 0.252

=====

Real				
	native.country	relationship	MAX(fnlwgt)	COUNT(*)
151	Mexico	Wife	412156	28
100	Hong	Wife	217921	7
191	Puerto-Rico	Own-child	538319	16
239	Yugoslavia	Unmarried	406518	1
75	Greece	Unmarried	227515	2

Synthetic				
	native.country	relationship	MAX(fnlwgt)	COUNT(*)
151	Mexico	Wife	199533.493530	4.000000
100	Hong	Wife	nan	nan
191	Puerto-Rico	Own-child	210829.953625	994.000000
239	Yugoslavia	Unmarried	194179.822551	38.000000
75	Greece	Unmarried	191611.068410	1.000000

SQL for Real:

SELECT `native.country`,relationship,MAX(fnlwgt), COUNT(*) FROM C1 GROUP BY `native.country`,relationship

Resulted in 241 records

SQL for Synthetic:

SELECT `native.country`,relationship,MAX(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY `native.country`,relationship

Resulted in 212 records

Normalized Euclidean distance for (fnlwgt): 14.25

Hellinger Distance: 0.399

=====

Real										
	sex	race	relationship	occupation	education	income	marital.status	native.country	SUM(fnlwgt)	COUNT(*)
12617	Male	White	Unmarried	Craft-repair	HS-grad	<=50K	Never-married	Puerto-Rico	nan	nan
9354	Male	White	Husband	Sales	Assoc-voc	<=50K	Married-civ-spouse	Columbia	nan	nan
12627	Male	White	Unmarried	Craft-repair	HS-grad	<=50K	Widowed	United-States	644187.000000	4.000000
8761	Male	White	Husband	Machine-op-inspct	11th	<=50K	Married-civ-spouse	Japan	nan	nan
382	Female	Black	Not-in-family	Adm-clerical	9th	<=50K	Divorced	United-States	nan	nan

Synthetic										
	sex	race	relationship	occupation	education	income	marital.status	native.country	SUM(fnlwgt)	COUNT(*)
12617	Male	White	Unmarried	Craft-repair	HS-grad	<=50K	Never-married	Puerto-Rico	344061.200592	2
9354	Male	White	Husband	Sales	Assoc-voc	<=50K	Married-civ-spouse	Columbia	710059.539570	4
12627	Male	White	Unmarried	Craft-repair	HS-grad	<=50K	Widowed	United-States	196908.784593	1
8761	Male	White	Husband	Machine-op-inspct	11th	<=50K	Married-civ-spouse	Japan	178170.547651	1
382	Female	Black	Not-in-family	Adm-clerical	9th	<=50K	Divorced	United-States	158023.412986	1

SQL for Real:

SELECT sex,race,relationship,occupation,education,income,`marital.status`,`native.country`,SUM(fnlwgt), COUNT(*) FROM C1 GROUP BY sex,race,relationship,occupation,education,income,`marital.status`,`native.country`

Resulted in 9390 records

SQL for Synthetic:

SELECT sex,race,relationship,occupation,education,income,`marital.status`,`native.country`,SUM(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY sex,race,relationship,occupation,education,income,`marital.status`,`native.country`

Resulted in 13048 records

Normalized Euclidean distance for (fnlwgt): 57.52

Hellinger Distance: 0.264

=====

Real				
	race	education	SUM(age)	COUNT(*)
42	Black	Doctorate	791	16
78	White	Prof-school	33990	748
5	Amer-Indian-Eskimo	7th-8th	530	10
54	Other	9th	426	15
11	Amer-Indian-Eskimo	HS-grad	6428	176

Synthetic				
	race	education	SUM(age)	COUNT(*)
42	Black	Doctorate	712.000000	15.000000
78	White	Prof-school	35718.000000	777.000000
5	Amer-Indian-Eskimo	7th-8th	282.000000	6.000000
54	Other	9th	338.000000	11.000000
11	Amer-Indian-Eskimo	HS-grad	2690.000000	73.000000

SQL for Real:

SELECT race,education,SUM(age), COUNT(*) FROM C1 GROUP BY race,education

Resulted in 80 records

SQL for Synthetic:

SELECT race,education,SUM(age), COUNT(*) FROM C1_syn_06 GROUP BY race,education

Resulted in 77 records

Normalized Euclidean distance for (age): 8.77

Hellinger Distance: 0.052

=====

Real								
	native.country	relationship	race	occupation	marital.status	sex	MIN(`hours.per.week`)	COUNT(*)
1978	Japan	Not-in-family	White	Other-service	Never-married	Female	nan	nan
545	Columbia	Not-in-family	Black	Adm-clerical	Never-married	Female	nan	nan
4058	Taiwan	Own-child	White	Adm-clerical	Never-married	Female	nan	nan
710	Columbia	Own-child	White	Prof-specialty	Married-civ-spouse	Male	nan	nan
4719	United-States	Other-relative	White	Exec-managerial	Married-civ-spouse	Male	40.000000	8.000000

Synthetic									
	native.country	relationship	race	occupation	marital.status	sex	MIN(`hours.per.week`)	COUNT(*)	
1978	Japan	Not-in-family	White	Other-service	Never-married	Female	39.954135	4	
545	Columbia	Not-in-family	Black	Adm-clerical	Never-married	Female	39.969043	3	
4058	Taiwan	Own-child	White	Adm-clerical	Never-married	Female	39.980069	2	
710	Columbia	Own-child	White	Prof-specialty	Married-civ-spouse	Male	40.023536	1	
4719	United-States	Other-relative	White	Exec-managerial	Married-civ-spouse	Male	39.972052	2	

SQL for Real:

SELECT `native.country`,relationship,race,occupation,`marital.status`,sex,MIN(`hours.per.week`), COUNT(*) FROM C1 GROUP BY `native.country`,relationship,race,occupation,`marital.status`,sex

Resulted in 3833 records

SQL for Synthetic:

SELECT `native.country`,relationship,race,occupation,`marital.status`,sex,MIN(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY `native.country`,relationship,race,occupation,`marital.status`,sex

Resulted in 5475 records

Normalized Euclidean distance for (`hours.per.week`): 40.22

Hellinger Distance: 0.325

=====

Real				
	marital.status	race	SUM(`hours.per.week`)	COUNT(*)
7	Married-AF-spouse	White	1293	33
31	Widowed	Other	337	9
29	Widowed	Asian-Pac-Islander	1370	39
5	Married-AF-spouse	Asian-Pac-Islander	60	1
30	Widowed	Black	6609	193

Synthetic				
	marital.status	race	SUM(`hours.per.week`)	COUNT(*)
7	Married-AF-spouse	White	279.452222	7.000000
31	Widowed	Other	359.069962	9.000000
29	Widowed	Asian-Pac-Islander	955.988301	24.000000
5	Married-AF-spouse	Asian-Pac-Islander	nan	nan
30	Widowed	Black	7283.116690	183.000000

SQL for Real:

SELECT `marital.status`,race,SUM(`hours.per.week`), COUNT(*) FROM C1 GROUP BY `marital.status`,race

Resulted in 33 records

SQL for Synthetic:

SELECT `marital.status`,race,SUM(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY `marital.status`,race

Resulted in 32 records

Normalized Euclidean distance for (`hours.per.week`): 5.66

Hellinger Distance: 0.046

=====

Real											
	workclass	race	native.country	education	marital.status	sex	occupation	relationship	income	SUM(age)	COUNT(*)
101	?	Black	Nicaragua	HS-grad	Never-married	Female	?	Own-child	<=50K	nan	nan
5583	Private	Black	United-States	HS-grad	Married-civ-spouse	Female	Machine-op-inspct	Own-child	<=50K	nan	nan
7826	Private	White	Hungary	Masters	Married-civ-spouse	Male	Prof-specialty	Husband	<=50K	nan	nan
2533	Local-gov	White	Ireland	Masters	Never-married	Female	Prof-specialty	Unmarried	<=50K	nan	nan
15890	Self-emp-not-inc	White	United-States	Some-college	Divorced	Female	Craft-repair	Not-in-family	<=50K	nan	nan

Synthetic											
	workclass	race	native.country	education	marital.status	sex	occupation	relationship	income	SUM(age)	COUNT(*)
101	?	Black	Nicaragua	HS-grad	Never-married	Female	?	Own-child	<=50K	20	1
5583	Private	Black	United-States	HS-grad	Married-civ-spouse	Female	Machine-op-inspct	Own-child	<=50K	37	1
7826	Private	White	Hungary	Masters	Married-civ-spouse	Male	Prof-specialty	Husband	<=50K	73	2
2533	Local-gov	White	Ireland	Masters	Never-married	Female	Prof-specialty	Unmarried	<=50K	31	1
15890	Self-emp-not-inc	White	United-States	Some-college	Divorced	Female	Craft-repair	Not-in-family	<=50K	41	1

SQL for Real:

SELECT workclass,race,`native.country`,education,`marital.status`,sex,occupation,relationship,income,SUM(age), COUNT(*) FROM C1 GROUP BY workclass,race,`native.country`,education,`marital.status`,sex,occupation,relationship,income

Resulted in 12664 records

SQL for Synthetic:

SELECT workclass,race,`native.country`,education,`marital.status`,sex,occupation,relationship,income,SUM(age), COUNT(*) FROM C1_syn_06 GROUP BY workclass,race,`native.country`,education,`marital.status`,sex,occupation,relationship,income

Resulted in 17127 records

Normalized Euclidean distance for (age): 64.61

Hellinger Distance: 0.252

=====

Real

	race	native.country	workclass	occupation	SUM(age)	COUNT(*)
985	White	?	Self-emp-inc	Other-service	nan	nan
245	Asian-Pac-Islander	Puerto-Rico	Private	Sales	nan	nan
749	Black	United-States	Federal-gov	Prof-specialty	579.000000	14.000000
935	Other	United-States	Private	Sales	406.000000	16.000000
481	Black	Haiti	Local-gov	Tech-support	nan	nan

Synthetic

	race	native.country	workclass	occupation	SUM(age)	COUNT(*)
985	White	?	Self-emp-inc	Other-service	77	2
245	Asian-Pac-Islander	Puerto-Rico	Private	Sales	524	11
749	Black	United-States	Federal-gov	Prof-specialty	191	4
935	Other	United-States	Private	Sales	647	25
481	Black	Haiti	Local-gov	Tech-support	23	1

SQL for Real:

SELECT race,`native.country`,workclass,occupation,SUM(age), COUNT(*) FROM C1 GROUP BY race,`native.country`,workclass,occupation

Resulted in 1475 records

SQL for Synthetic:

SELECT race,`native.country`,workclass,occupation,SUM(age), COUNT(*) FROM C1_syn_06 GROUP BY race,`native.country`,workclass,occupation

Resulted in 1993 records

Normalized Euclidean distance for (age): 28.14

Hellinger Distance: 0.363

=====

Real

	race	sex	occupation	AVG(fnlwgt)	COUNT(*)
121	White	Female	Sales	185839.636087	1646.000000
70	Black	Male	Adm-clerical	234904.921466	191.000000
37	Asian-Pac-Islander	Female	Protective-serv	190208.000000	4.000000
99	Other	Male	Exec-managerial	205629.153846	13.000000
50	Asian-Pac-Islander	Male	Protective-serv	136641.235294	17.000000

Synthetic

	race	sex	occupation	AVG(fnlwgt)	COUNT(*)
121	White	Female	Sales	181866.219619	1802
70	Black	Male	Adm-clerical	179218.295175	157
37	Asian-Pac-Islander	Female	Protective-serv	181613.649040	1
99	Other	Male	Exec-managerial	182517.191619	10
50	Asian-Pac-Islander	Male	Protective-serv	174580.186010	11

SQL for Real:

SELECT race,sex,occupation,AVG(fnlwgt), COUNT(*) FROM C1 GROUP BY race,sex,occupation

Resulted in 137 records

SQL for Synthetic:

SELECT race,sex,occupation,AVG(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY race,sex,occupation

Resulted in 139 records

Normalized Euclidean distance for (fnlwgt): 11.58

Hellinger Distance: 0.056

=====

Real

	education	sex	workclass	marital.status	native.country	race	relationship	income	MAX(capital)	COUNT(*)
8012	Some-college	Female	Private	Married-civ-spouse	Jamaica	White	Wife	<=50K	nan	nan
23	10th	Female	?	Never-married	United-States	Black	Other-relative	<=50K	nan	nan
3024	Assoc-voc	Male	Private	Never-married	United-States	White	Own-child	<=50K	4416.000000	51.000000
9045	Some-college	Male	Self-emp-not-inc	Divorced	United-States	White	Own-child	<=50K	0.000000	3.000000
5732	HS-grad	Male	?	Never-married	Columbia	White	Not-in-family	<=50K	nan	nan

Synthetic

	education	sex	workclass	marital.status	native.country	race	relationship	income	MAX(capital)	COUNT(*)
8012	Some-college	Female	Private	Married-civ-spouse	Jamaica	White	Wife	<=50K	26.496358	1
23	10th	Female	?	Never-married	United-States	Black	Other-relative	<=50K	-79.871959	1
3024	Assoc-voc	Male	Private	Never-married	United-States	White	Own-child	<=50K	254.910793	70
9045	Some-college	Male	Self-emp-not-inc	Divorced	United-States	White	Own-child	<=50K	175.276843	3
5732	HS-grad	Male	?	Never-married	Columbia	White	Not-in-family	<=50K	-69.984767	1

SQL for Real:

SELECT education,sex,workclass,`marital.status`,`native.country`,race,relationship,income,MAX(capital), COUNT(*) FROM C1 GROUP BY education,sex,workclass,`marital.status`,`native.country`,race,relationship,income

Resulted in 6863 records

SQL for Synthetic:

SELECT education,sex,workclass,`marital.status`,`native.country`,race,relationship,income,MAX(capital), COUNT(*) FROM C1_syn_06 GROUP BY education,sex,workclass,`marital.status`,`native.country`,race,relationship,income

Resulted in 9165 records

Normalized Euclidean distance for (capital): 50.11

Hellinger Distance: 0.311

=====

Real				
	workclass	marital.status	MIN(age)	COUNT(*)
31	Self-emp-inc	Separated	26	25
32	Self-emp-inc	Widowed	33	41
37	Self-emp-not-inc	Never-married	17	613
9	Federal-gov	Married-civ-spouse	22	721
0	?	Divorced	20	270

Synthetic				
	workclass	marital.status	MIN(age)	COUNT(*)
31	Self-emp-inc	Separated	26.000000	20.000000
32	Self-emp-inc	Widowed	25.000000	44.000000
37	Self-emp-not-inc	Never-married	17.000000	635.000000
9	Federal-gov	Married-civ-spouse	23.000000	770.000000
0	?	Divorced	19.000000	304.000000

SQL for Real:
SELECT workclass,`marital.status`,MIN(age), COUNT(*) FROM C1 GROUP BY workclass,`marital.status`
Resulted in 47 records

SQL for Synthetic:
SELECT workclass,`marital.status`,MIN(age), COUNT(*) FROM C1_syn_06 GROUP BY workclass,`marital.status`
Resulted in 47 records

Normalized Euclidean distance for (age): 6.86

Hellinger Distance: 0.018

=====

Real						
	education	sex	native.country	race	AVG(age)	COUNT(*)
369	7th-8th	Female	United-States	Black	54.783784	37.000000
263	1st-4th	Male	Cambodia	Asian-Pac-Islander	37.000000	1.000000
663	Bachelors	Female	Columbia	Asian-Pac-Islander	nan	nan
270	1st-4th	Male	Laos	White	nan	nan
781	Bachelors	Male	Portugal	White	46.000000	1.000000

Synthetic						
	education	sex	native.country	race	AVG(age)	COUNT(*)
369	7th-8th	Female	United-States	Black	56.214286	14
263	1st-4th	Male	Cambodia	Asian-Pac-Islander	37.000000	1
663	Bachelors	Female	Columbia	Asian-Pac-Islander	43.250000	4
270	1st-4th	Male	Laos	White	26.000000	1
781	Bachelors	Male	Portugal	White	34.000000	1

SQL for Real:
SELECT education,sex,`native.country`,race,AVG(age), COUNT(*) FROM C1 GROUP BY education,sex,`native.country`,race
Resulted in 1171 records

SQL for Synthetic:
SELECT education,sex,`native.country`,race,AVG(age), COUNT(*) FROM C1_syn_06 GROUP BY education,sex,`native.country`,race
Resulted in 1368 records

Normalized Euclidean distance for (age): 24.76

Hellinger Distance: 0.378

=====

Real		
	income	MAX(fnlwgt) COUNT(*)
0	<=50K	1490400 37155
1	>50K	1226583 11687

Synthetic		
	income	MAX(fnlwgt) COUNT(*)
0	<=50K	1461813.868173 44080.000000
1	>50K	1362953.001310 4762.000000

SQL for Real:
SELECT income,MAX(fnlwgt), COUNT(*) FROM C1 GROUP BY income
Resulted in 2 records

SQL for Synthetic:
SELECT income,MAX(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY income
Resulted in 2 records

Normalized Euclidean distance for (fnlwgt): 1.41

Hellinger Distance: 0.137

=====

Real										
	native.country	income	occupation	relationship	sex	education	marital.status	workclass	SUM(capital)	COUNT(*)
9034	Scotland	<=50K	Other-service	Unmarried	Female	10th	Divorced	Private	nan	nan
10359	United-States	<=50K	Craft-repair	Not-in-family	Male	7th-8th	Divorced	Local-gov	nan	nan
2540	Hungary	<=50K	Craft-repair	Husband	Male	10th	Married-civ-spouse	Private	nan	nan
7539	Puerto-Rico	<=50K	Other-service	Unmarried	Male	12th	Never-married	State-gov	nan	nan
12454	United-States	<=50K	Prof-specialty	Husband	Male	HS-grad	Married-civ-spouse	State-gov	0.000000	3.000000

Synthetic										
	native.country	income	occupation	relationship	sex	education	marital.status	workclass	SUM(capital)	COUNT(*)
9034	Scotland	<=50K	Other-service	Unmarried	Female	10th	Divorced	Private	-24.599306	1
10359	United-States	<=50K	Craft-repair	Not-in-family	Male	7th-8th	Divorced	Local-gov	89.013609	1
2540	Hungary	<=50K	Craft-repair	Husband	Male	10th	Married-civ-spouse	Private	76.415721	4
7539	Puerto-Rico	<=50K	Other-service	Unmarried	Male	12th	Never-married	State-gov	-132.054908	1
12454	United-States	<=50K	Prof-specialty	Husband	Male	HS-grad	Married-civ-spouse	State-gov	-296.496987	3

SQL for Real:
SELECT `native.country`,income,occupation,relationship,sex,education,`marital.status`,workclass,SUM(capital), COUNT(*) FROM C1 GROUP BY `native.country`,income,occupation,relationship,sex,education,`marital.status`,workclass

Resulted in 10400 records

SQL for Synthetic:
SELECT `native.country`,income,occupation,relationship,sex,education,`marital.status`,workclass,SUM(capital), COUNT(*) FROM C1_syn_06 GROUP BY `native.country`,income,occupation,relationship,sex,education,`marital.status`,workclass

Resulted in 14639 records

Normalized Euclidean distance for (capital): 62.27

Hellinger Distance: 0.26

=====

Real						
	marital.status	income	workclass	sex	AVG(capital)	COUNT(*)
70	Married-spouse-absent	<=50K	?	Male	-90.800000	20
90	Married-spouse-absent	>50K	Self-emp-inc	Male	0.000000	1
115	Never-married	>50K	Private	Female	5629.491124	169
134	Separated	<=50K	Self-emp-not-inc	Male	-51.307692	52
31	Married-AF-spouse	<=50K	Private	Female	0.000000	8

Synthetic						
	marital.status	income	workclass	sex	AVG(capital)	COUNT(*)
70	Married-spouse-absent	<=50K	?	Male	0.889995	15.000000
90	Married-spouse-absent	>50K	Self-emp-inc	Male	-55.292177	1.000000
115	Never-married	>50K	Private	Female	-19.406831	48.000000
134	Separated	<=50K	Self-emp-not-inc	Male	19.938936	51.000000
31	Married-AF-spouse	<=50K	Private	Female	33.555565	3.000000

SQL for Real:
SELECT `marital.status`,income,workclass,sex,AVG(capital), COUNT(*) FROM C1 GROUP BY `marital.status`,income,workclass,sex

Resulted in 175 records

SQL for Synthetic:
SELECT `marital.status`,income,workclass,sex,AVG(capital), COUNT(*) FROM C1_syn_06 GROUP BY `marital.status`,income,workclass,sex

Resulted in 153 records

Normalized Euclidean distance for (capital): 12.25

Hellinger Distance: 0.151

=====

Real						
	income	native.country	relationship	race	education	SUM(`hours.per.week`) COUNT(*)
391	<=50K	Dominican-Republic	Husband	Asian-Pac-Islander	HS-grad	nan nan
1794	<=50K	Scotland	Husband	Black	11th	nan nan
1361	<=50K	Peru	Husband	White	1st-4th	nan nan
185	<=50K	Canada	Other-relative	White	Bachelors	nan nan
1977	<=50K	Thailand	Unmarried	Asian-Pac-Islander	Bachelors	60.000000 1.000000

Synthetic						
	income	native.country	relationship	race	education	SUM(`hours.per.week`) COUNT(*)
391	<=50K	Dominican-Republic	Husband	Asian-Pac-Islander	HS-grad	40.023817 1
1794	<=50K	Scotland	Husband	Black	11th	79.914611 2
1361	<=50K	Peru	Husband	White	1st-4th	79.879051 2
185	<=50K	Canada	Other-relative	White	Bachelors	39.995473 1
1977	<=50K	Thailand	Unmarried	Asian-Pac-Islander	Bachelors	80.024028 2

SQL for Real:
SELECT income,`native.country`,relationship,race,education,SUM(`hours.per.week`), COUNT(*) FROM C1 GROUP BY income,`native.country`,relationship,race,education

Resulted in 2408 records

SQL for Synthetic:
SELECT income,`native.country`,relationship,race,education,SUM(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY income,`native.country`,relationship,race,education

Resulted in 2863 records

Normalized Euclidean distance for (`hours.per.week`): 32.4

Hellinger Distance: 0.369

=====

Real			
	occupation	MAX(capital)	COUNT(*)
4	Exec-managerial	99999	6086
9	Priv-house-serv	25236	242
6	Handlers-cleaners	99999	2072
0	?	99999	2809
3	Craft-repair	99999	6112

Synthetic			
	occupation	MAX(capital)	COUNT(*)
4	Exec-managerial	628.300592	6095.000000
9	Priv-house-serv	292.403797	223.000000
6	Handlers-cleaners	523.214732	2129.000000
0	?	483.749628	2840.000000
3	Craft-repair	371.249158	5926.000000

SQL for Real:
SELECT occupation,MAX(capital), COUNT(*) FROM C1 GROUP BY occupation
Resulted in 15 records

SQL for Synthetic:
SELECT occupation,MAX(capital), COUNT(*) FROM C1_syn_06 GROUP BY occupation
Resulted in 15 records

Normalized Euclidean distance for (capital): 3.87

Hellinger Distance: 0.027

=====

Real							
	income	native.country	occupation	education	relationship	marital.status	AVG(capital) COUNT(*)
344	<=50K	?	Prof-specialty	Masters	Not-in-family	Separated	0.000000 1.000000
942	<=50K	Columbia	Other-service	Assoc-voc	Husband	Married-civ-spouse	nan nan
5310	<=50K	Puerto-Rico	Sales	HS-grad	Other-relative	Widowed	nan nan
8994	>50K	United-States	Prof-specialty	Doctorate	Not-in-family	Divorced	3019.818182 11.000000
6925	<=50K	United-States	Handlers-cleaners	Some-college	Husband	Married-civ-spouse	195.215190 79.000000

Synthetic							
	income	native.country	occupation	education	relationship	marital.status	AVG(capital) COUNT(*)
344	<=50K	?	Prof-specialty	Masters	Not-in-family	Separated	153.970754 1
942	<=50K	Columbia	Other-service	Assoc-voc	Husband	Married-civ-spouse	8.134245 3
5310	<=50K	Puerto-Rico	Sales	HS-grad	Other-relative	Widowed	116.307822 1
8994	>50K	United-States	Prof-specialty	Doctorate	Not-in-family	Divorced	-83.057868 4
6925	<=50K	United-States	Handlers-cleaners	Some-college	Husband	Married-civ-spouse	0.953003 58

SQL for Real:
SELECT income,`native.country`,occupation,education,relationship,`marital.status`,AVG(capital), COUNT(*) FROM C1 GROUP BY income,`native.country`,occupation,education,relationship,`marital.status`
Resulted in 6304 records

SQL for Synthetic:
SELECT income,`native.country`,occupation,education,relationship,`marital.status`,AVG(capital), COUNT(*) FROM C1_syn_06 GROUP BY income,`native.country`,occupation,education,relationship,`marital.status`
Resulted in 9107 records

Normalized Euclidean distance for (capital): 50.49

Hellinger Distance: 0.281

=====

Real										
	workclass	marital.status	race	income	education	native.country	sex	occupation	relationship	MIN(capital) COUNT(*)
5081	Private	Divorced	White	<=50K	HS-grad	United-States	Male	Sales	Not-in-family	-1669.000000 37.000000
10933	Private	Never-married	White	<=50K	Bachelors	United-States	Male	?	Not-in-family	nan nan
15493	Self-emp-not-inc	Never-married	White	<=50K	9th	Puerto-Rico	Male	Exec-managerial	Unmarried	nan nan
12826	Private	Separated	White	<=50K	HS-grad	Nicaragua	Female	Adm-clerical	Not-in-family	nan nan
12446	Private	Separated	Black	<=50K	HS-grad	Puerto-Rico	Female	Machine-op-inspct	Not-in-family	nan nan

Synthetic										
	workclass	marital.status	race	income	education	native.country	sex	occupation	relationship	MIN(capital) COUNT(*)
5081	Private	Divorced	White	<=50K	HS-grad	United-States	Male	Sales	Not-in-family	-104.632923 14
10933	Private	Never-married	White	<=50K	Bachelors	United-States	Male	?	Not-in-family	-44.890722 1
15493	Self-emp-not-inc	Never-married	White	<=50K	9th	Puerto-Rico	Male	Exec-managerial	Unmarried	-117.521361 1
12826	Private	Separated	White	<=50K	HS-grad	Nicaragua	Female	Adm-clerical	Not-in-family	-228.399526 1
12446	Private	Separated	Black	<=50K	HS-grad	Puerto-Rico	Female	Machine-op-inspct	Not-in-family	25.810966 1

SQL for Real:
SELECT workclass,`marital.status`,race,income,education,`native.country`,sex,occupation,relationship,MIN(capital), COUNT(*) FROM C1 GROUP BY workclass,`marital.status`,race,income,education,`native.country`,sex,occupation,relationship
Resulted in 12664 records

SQL for Synthetic:
SELECT workclass,`marital.status`,race,income,education,`native.country`,sex,occupation,relationship,MIN(capital), COUNT(*) FROM C1_syn_06 GROUP BY workclass,`marital.status`,race,income,education,`native.country`,sex,occupation,relationship
Resulted in 17127 records

Normalized Euclidean distance for (capital): 64.61

Hellinger Distance: 0.252

=====

	occupation	sex	relationship	marital.status	native.country	workclass	education	income	race	SUM(`hours.per.week`)	COUNT(*)
17073	Transport-moving	Male	Unmarried	Divorced	Puerto-Rico	Private	Some-college	<=50K	White	nan	nan
341	?	Female	Unmarried	Divorced	United-States	?	Assoc-acdm	<=50K	White	1.000000	1.000000
9876	Other-service	Female	Unmarried	Divorced	Germany	Private	Some-college	<=50K	White	nan	nan
273	?	Female	Own-child	Never-married	Puerto-Rico	?	Some-college	<=50K	White	nan	nan
2618	Adm-clerical	Male	Husband	Married-civ-spouse	United-States	Local-gov	Assoc-acdm	<=50K	Black	nan	nan

Synthetic

	occupation	sex	relationship	marital.status	native.country	workclass	education	income	race	SUM(`hours.per.week`)	COUNT(*)
17073	Transport-moving	Male	Unmarried	Divorced	Puerto-Rico	Private	Some-college	<=50K	White	39.999859	1
341	?	Female	Unmarried	Divorced	United-States	?	Assoc-acdm	<=50K	White	80.085244	2
9876	Other-service	Female	Unmarried	Divorced	Germany	Private	Some-college	<=50K	White	40.040348	1
273	?	Female	Own-child	Never-married	Puerto-Rico	?	Some-college	<=50K	White	471.532390	12
2618	Adm-clerical	Male	Husband	Married-civ-spouse	United-States	Local-gov	Assoc-acdm	<=50K	Black	39.983237	1

SQL for Real:
SELECT occupation,sex,relationship,`marital.status`,`native.country`,workclass,education,income,race,SUM(`hours.per.week`), COUNT(*) FROM C1 GROUP BY occupation,sex,relationship,`marital.status`,`native.country`,workclass,education,income,race
Resulted in 12664 records

SQL for Synthetic:
SELECT occupation,sex,relationship,`marital.status`,`native.country`,workclass,education,income,race,SUM(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY occupation,sex,relationship,`marital.status`,`native.country`,workclass,education,income,race
Resulted in 17127 records

Normalized Euclidean distance for (`hours.per.week`): 64.61

Hellinger Distance: 0.252

=====

	sex	education	occupation	workclass	marital.status	MAX(`hours.per.week`)	COUNT(*)				
3178	Male	HS-grad	Sales	Self-emp-inc	Widowed	45.000000	2.000000				
47	Female	10th	Handlers-cleaners	Private	Never-married	34.000000	3.000000				
3295	Male	Masters	Handlers-cleaners	Private	Married-civ-spouse	46.000000	1.000000				
1377	Female	Prof-school	Exec-managerial	Private	Never-married	40.000000	1.000000				
1540	Female	Some-college	Other-service	Local-gov	Divorced	45.000000	9.000000				

Synthetic

	sex	education	occupation	workclass	marital.status	MAX(`hours.per.week`)	COUNT(*)				
3178	Male	HS-grad	Sales	Self-emp-inc	Widowed	39.925747	1				
47	Female	10th	Handlers-cleaners	Private	Never-married	40.020518	5				
3295	Male	Masters	Handlers-cleaners	Private	Married-civ-spouse	40.019872	4				
1377	Female	Prof-school	Exec-managerial	Private	Never-married	40.009540	1				
1540	Female	Some-college	Other-service	Local-gov	Divorced	40.032732	6				

SQL for Real:
SELECT sex,education,occupation,workclass,`marital.status`,MAX(`hours.per.week`), COUNT(*) FROM C1 GROUP BY sex,education,occupation,workclass,`marital.status`
Resulted in 3644 records

SQL for Synthetic:
SELECT sex,education,occupation,workclass,`marital.status`,MAX(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY sex,education,occupation,workclass,`marital.status`
Resulted in 3741 records

Normalized Euclidean distance for (`hours.per.week`): 51.15

Hellinger Distance: 0.105

=====

	native.country	income	workclass	sex	occupation	relationship	marital.status	education	SUM(capital)	COUNT(*)	
8367	Puerto-Rico	<=50K	State-gov	Male	Adm-clerical	Own-child	Never-married	HS-grad	nan	nan	
11056	United-States	<=50K	Private	Female	Other-service	Not-in-family	Divorced	5th-6th	0.000000	2.000000	
3094	India	<=50K	Self-emp-not-inc	Female	Farming-fishing	Wife	Never-married	Assoc-voc	nan	nan	
12074	United-States	<=50K	Private	Male	Machine-op-inspct	Husband	Never-married	HS-grad	nan	nan	
10588	United-States	<=50K	Private	Female	Adm-clerical	Not-in-family	Never-married	12th	0.000000	2.000000	

Synthetic										
	native.country	income	workclass	sex	occupation	relationship	marital.status	education	SUM(capital)	COUNT(*)
8367	Puerto-Rico	<=50K	State-gov	Male	Adm-clerical	Own-child	Never-married	HS-grad	-29.688122	1
11056	United-States	<=50K	Private	Female	Other-service	Not-in-family	Divorced	5th-6th	-48.462160	1
3094	India	<=50K	Self-emp-not-inc	Female	Farming-fishing	Wife	Never-married	Assoc-voc	33.565662	1
12074	United-States	<=50K	Private	Male	Machine-op-inspct	Husband	Never-married	HS-grad	-80.726735	1
10588	United-States	<=50K	Private	Female	Adm-clerical	Not-in-family	Never-married	12th	316.195616	5

SQL for Real:
SELECT `native.country`,income,workclass,sex,occupation,relationship,`marital.status`,education,SUM(capital), COUNT(*) FROM C1 GROUP BY `native.country`,income,workclass,sex,occupation,relationship,`marital.status`,education

Resulted in 10400 records

SQL for Synthetic:
SELECT `native.country`,income,workclass,sex,occupation,relationship,`marital.status`,education,SUM(capital), COUNT(*) FROM C1_syn_06 GROUP BY `native.country`,income,workclass,sex,occupation,relationship,`marital.status`,education

Resulted in 14639 records

Normalized Euclidean distance for (capital): 62.27

Hellinger Distance: 0.26

=====

Real			
	sex	native.country	SUM(`hours.per.week`) COUNT(*)
41	Female	Yugoslavia	156 5
36	Female	Taiwan	646 19
38	Female	Trinidad&Tobago	487 14
59	Male	Hungary	498 12
5	Female	Cuba	1802 50

Synthetic			
	sex	native.country	SUM(`hours.per.week`) COUNT(*)
41	Female	Yugoslavia	3280.454909 82.000000
36	Female	Taiwan	2551.638069 64.000000
38	Female	Trinidad&Tobago	119.059441 3.000000
59	Male	Hungary	29312.906923 731.000000
5	Female	Cuba	119.971744 3.000000

SQL for Real:
SELECT sex,`native.country`,SUM(`hours.per.week`), COUNT(*) FROM C1 GROUP BY sex,`native.country`
Resulted in 83 records

SQL for Synthetic:
SELECT sex,`native.country`,SUM(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY sex,`native.country`
Resulted in 79 records

Normalized Euclidean distance for (`hours.per.week`): 8.83

Hellinger Distance: 0.399

=====

Real				
	marital.status	education	AVG(age)	COUNT(*)
97	Widowed	Bachelors	56.472000	125
11	Divorced	HS-grad	42.537666	2416
15	Divorced	Some-college	42.214103	1560
7	Divorced	Assoc-acdm	41.628571	280
12	Divorced	Masters	46.871935	367

Synthetic				
	marital.status	education	AVG(age)	COUNT(*)
97	Widowed	Bachelors	56.913386	127.000000
11	Divorced	HS-grad	42.244996	2498.000000
15	Divorced	Some-college	42.399467	1502.000000
7	Divorced	Assoc-acdm	40.063830	282.000000
12	Divorced	Masters	46.022556	399.000000

SQL for Real:
SELECT `marital.status`,education,AVG(age), COUNT(*) FROM C1 GROUP BY `marital.status`,education
Resulted in 104 records

SQL for Synthetic:
SELECT `marital.status`,education,AVG(age), COUNT(*) FROM C1_syn_06 GROUP BY `marital.status`,education
Resulted in 101 records

Normalized Euclidean distance for (age): 10.0

Hellinger Distance: 0.034

=====

Real			
	relationship	SUM(`hours.per.week`)	COUNT(*)
5	Wife	85617	2331
1	Not-in-family	509987	12583
3	Own-child	251253	7581
2	Other-relative	55915	1506
4	Unmarried	200759	5125

Synthetic			
	relationship	SUM(`hours.per.week`)	COUNT(*)
5	Wife	99106.834118	2479.000000
1	Not-in-family	497260.007983	12433.000000
3	Own-child	326930.781923	8227.000000
2	Other-relative	53879.800868	1351.000000
4	Unmarried	199780.219536	4999.000000

SQL for Real:
SELECT relationship,SUM(`hours.per.week`), COUNT(*) FROM C1 GROUP BY relationship
Resulted in 6 records

SQL for Synthetic:
SELECT relationship,SUM(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY relationship
Resulted in 6 records

Normalized Euclidean distance for (`hours.per.week`): 2.45

Hellinger Distance: 0.015

=====

Real							
	sex	relationship	native.country	marital.status	workclass	MAX(capital)	COUNT(*)
796	Female	Unmarried	Ireland	Never-married	Local-gov	nan	nan
1066	Female	Wife	Philippines	Married-civ-spouse	Self-emp-not-inc	nan	nan
522	Female	Own-child	Haiti	Never-married	Local-gov	0.000000	1.000000
771	Female	Unmarried	Holand-Netherlands	Never-married	Local-gov	nan	nan
1433	Male	Not-in-family	Laos	Separated	Private	nan	nan

Synthetic							
	sex	relationship	native.country	marital.status	workclass	MAX(capital)	COUNT(*)
796	Female	Unmarried	Ireland	Never-married	Local-gov	113.084300	3
1066	Female	Wife	Philippines	Married-civ-spouse	Self-emp-not-inc	285.861586	2
522	Female	Own-child	Haiti	Never-married	Local-gov	162.814864	4
771	Female	Unmarried	Holand-Netherlands	Never-married	Local-gov	-58.475034	1
1433	Male	Not-in-family	Laos	Separated	Private	127.435551	5

SQL for Real:
SELECT sex,relationship,`native.country`,`marital.status`,workclass,MAX(capital), COUNT(*) FROM C1 GROUP BY sex,relationship,`native.country`,`marital.status`,workclass
Resulted in 1609 records

SQL for Synthetic:
SELECT sex,relationship,`native.country`,`marital.status`,workclass,MAX(capital), COUNT(*) FROM C1_syn_06 GROUP BY sex,relationship,`native.country`,`marital.status`,workclass
Resulted in 2047 records

Normalized Euclidean distance for (capital): 28.3

Hellinger Distance: 0.373

=====

Real											
	sex	native.country	education	income	marital.status	workclass	relationship	race	occupation	AVG(age)	COUNT(*)
13628	Male	United-States	10th	>50K	Married-civ-spouse	Private	Husband	White	Prof-specialty	nan	nan
8521	Male	Columbia	HS-grad	<=50K	Never-married	Private	Not-in-family	White	Transport-moving	nan	nan
16613	Male	United-States	Some-college	<=50K	Married-civ-spouse	Local-gov	Husband	White	Prof-specialty	38.800000	5.000000
3936	Female	Puerto-Rico	Masters	<=50K	Divorced	Private	Not-in-family	White	Tech-support	nan	nan
159	Female	?	HS-grad	<=50K	Divorced	Private	Not-in-family	White	Adm-clerical	nan	nan

Synthetic											
	sex	native.country	education	income	marital.status	workclass	relationship	race	occupation	AVG(age)	COUNT(*)
13628	Male	United-States	10th	>50K	Married-civ-spouse	Private	Husband	White	Prof-specialty	59.000000	1
8521	Male	Columbia	HS-grad	<=50K	Never-married	Private	Not-in-family	White	Transport-moving	40.750000	8
16613	Male	United-States	Some-college	<=50K	Married-civ-spouse	Local-gov	Husband	White	Prof-specialty	51.333333	3
3936	Female	Puerto-Rico	Masters	<=50K	Divorced	Private	Not-in-family	White	Tech-support	50.000000	1
159	Female	?	HS-grad	<=50K	Divorced	Private	Not-in-family	White	Adm-clerical	29.000000	2

SQL for Real:
SELECT sex,`native.country`,education,income,`marital.status`,workclass,relationship,race,occupation,AVG(age), COUNT(*) FROM C1 GROUP BY sex,`native.country`,education,income,`marital.status`,workclass,relationship,race,occupation
Resulted in 12664 records

SQL for Synthetic:
SELECT sex,`native.country`,education,income,`marital.status`,workclass,relationship,race,occupation,AVG(age), COUNT(*) FROM C1_syn_06 GROUP BY sex,`native.country`,education,income,`marital.status`,workclass,relationship,race,occupation
Resulted in 17127 records

Normalized Euclidean distance for (age): 64.61

Hellinger Distance: 0.252

=====

Real				
	sex	native.country	income	MIN(fnlwgt) COUNT(*)
116	Male	Iran	<=50K	59469 27
108	Male	Honduras	<=50K	149640 8
29	Female	Honduras	>50K	51835 1
63	Female	Scotland	<=50K	89813 7
153	Male	Trinidad&Tobago	>50K	154863 2

Synthetic				
	sex	native.country	income	MIN(fnlwgt) COUNT(*)
116	Male	Iran	<=50K	179545.975184 146.000000
108	Male	Honduras	<=50K	nan nan
29	Female	Honduras	>50K	nan nan
63	Female	Scotland	<=50K	179547.146430 232.000000
153	Male	Trinidad&Tobago	>50K	nan nan

SQL for Real:
SELECT sex,`native.country`,income,MIN(fnlwgt), COUNT(*) FROM C1 GROUP BY sex,`native.country`,income
Resulted in 160 records

SQL for Synthetic:
SELECT sex,`native.country`,income,MIN(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY sex,`native.country`,income
Resulted in 130 records

Normalized Euclidean distance for (fnlwgt): 11.18

Hellinger Distance: 0.415

=====

Real								
	occupation	relationship	marital.status	education	race	native.country	MIN(capital)	COUNT(*)
8376	Protective-serv	Not-in-family	Never-married	Bachelors	White	United-States	-1876.000000	18.000000
9554	Tech-support	Husband	Married-civ-spouse	Assoc-acdm	White	Ecuador	nan	nan
2291	Craft-repair	Husband	Married-civ-spouse	Some-college	Black	Columbia	nan	nan
4545	Handlers-cleaners	Not-in-family	Never-married	12th	White	Columbia	nan	nan
9917	Tech-support	Unmarried	Divorced	HS-grad	White	United-States	0.000000	23.000000

Synthetic								
	occupation	relationship	marital.status	education	race	native.country	MIN(capital)	COUNT(*)
8376	Protective-serv	Not-in-family	Never-married	Bachelors	White	United-States	-195.937056	8
9554	Tech-support	Husband	Married-civ-spouse	Assoc-acdm	White	Ecuador	-132.986368	2
2291	Craft-repair	Husband	Married-civ-spouse	Some-college	Black	Columbia	-13.495988	1
4545	Handlers-cleaners	Not-in-family	Never-married	12th	White	Columbia	108.969914	1
9917	Tech-support	Unmarried	Divorced	HS-grad	White	United-States	-4.834948	5

SQL for Real:
SELECT occupation,relationship,`marital.status`,education,race,`native.country`,MIN(capital), COUNT(*) FROM C1 GROUP BY occupation,relationship,`marital.status`,education,race,`native.country`
Resulted in 7243 records

SQL for Synthetic:
SELECT occupation,relationship,`marital.status`,education,race,`native.country`,MIN(capital), COUNT(*) FROM C1_syn_06 GROUP BY occupation,relationship,`marital.status`,education,race,`native.country`
Resulted in 10524 records

Normalized Euclidean distance for (capital): 52.81

Hellinger Distance: 0.251

=====

Real				
	income	marital.status	occupation	MAX(fnlwgt) COUNT(*)
26	<=50K	Married-civ-spouse	Craft-repair	864960 2566
105	>50K	Divorced	Sales	456236 87
128	>50K	Married-civ-spouse	Tech-support	617021 356
82	<=50K	Widowed	Adm-clerical	953588 238
155	>50K	Separated	Craft-repair	352045 9

Synthetic				
	income	marital.status	occupation	MAX(fnlwgt) COUNT(*)
26	<=50K	Married-civ-spouse	Craft-repair	1461813.868173 3208.000000
105	>50K	Divorced	Sales	184911.649640 16.000000
128	>50K	Married-civ-spouse	Tech-support	196184.321460 185.000000
82	<=50K	Widowed	Adm-clerical	1340361.588728 283.000000
155	>50K	Separated	Craft-repair	182946.894151 6.000000

SQL for Real:
SELECT income,`marital.status`,occupation,MAX(fnlwgt), COUNT(*) FROM C1 GROUP BY income,`marital.status`,occupation
Resulted in 179 records

SQL for Synthetic:
SELECT income,`marital.status`,occupation,MAX(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY income,`marital.status`,occupation
Resulted in 159 records

Normalized Euclidean distance for (fnlwgt): 12.25

Hellinger Distance: 0.159

=====

Real					
	race	education	income	native.country	MIN(fnlwgt) COUNT(*)
885	White	Bachelors	<=50K	Greece	96062.000000 2.000000
861	White	Assoc-voc	<=50K	United-States	13769.000000 1243.000000
868	White	Assoc-voc	>50K	Ireland	nan nan
269	Black	11th	<=50K	Japan	nan nan
661	White	11th	>50K	?	118474.000000 1.000000

Synthetic					
	race	education	income	native.country	MIN(fnlwgt) COUNT(*)
885	White	Bachelors	<=50K	Greece	184431.189179 1
861	White	Assoc-voc	<=50K	United-States	133670.530895 906
868	White	Assoc-voc	>50K	Ireland	178494.159521 3
269	Black	11th	<=50K	Japan	177913.690013 3
661	White	11th	>50K	?	165472.155317 1

SQL for Real:
SELECT race,education,income,`native.country`,MIN(fnlwgt), COUNT(*) FROM C1 GROUP BY race,education,income,`native.country`
Resulted in 1070 records

SQL for Synthetic:
SELECT race,education,income,`native.country`,MIN(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY race,education,income,`native.country`
Resulted in 1158 records

Normalized Euclidean distance for (fnlwgt): 23.52

Hellinger Distance: 0.4

=====

Real				
	race	workclass	SUM(capital)	COUNT(*)
4	Amer-Indian-Eskimo	Self-emp-inc	17909	2
28	White	?	1042877	2337
9	Asian-Pac-Islander	Local-gov	113856	59
19	Black	Self-emp-not-inc	236814	136
22	Other	Federal-gov	1506	11

Synthetic				
	race	workclass	SUM(capital)	COUNT(*)
4	Amer-Indian-Eskimo	Self-emp-inc	nan	nan
28	White	?	1797.469440	2460.000000
9	Asian-Pac-Islander	Local-gov	1368.861830	50.000000
19	Black	Self-emp-not-inc	1065.757357	93.000000
22	Other	Federal-gov	190.408563	6.000000

SQL for Real:
SELECT race,workclass,SUM(capital), COUNT(*) FROM C1 GROUP BY race,workclass
Resulted in 35 records

SQL for Synthetic:
SELECT race,workclass,SUM(capital), COUNT(*) FROM C1_syn_06 GROUP BY race,workclass
Resulted in 34 records

Normalized Euclidean distance for (capital): 5.83

Hellinger Distance: 0.044

=====

Real						
	occupation	workclass	relationship	race	MIN(fnlwgt)	COUNT(*)
652	Other-service	Private	Wife	Black	24364	30
393	Farming-fishing	Local-gov	Not-in-family	White	70655	5
839	Prof-specialty	State-gov	Own-child	Black	122660	3
1128	Transport-moving	Private	Unmarried	White	22422	83
700	Other-service	State-gov	Unmarried	Black	73296	13

Synthetic						
	occupation	workclass	relationship	race	MIN(fnlwgt)	COUNT(*)
652	Other-service	Private	Wife	Black	162351.735777	31.000000
393	Farming-fishing	Local-gov	Not-in-family	White	179009.244679	7.000000
839	Prof-specialty	State-gov	Own-child	Black	171586.887237	4.000000
1128	Transport-moving	Private	Unmarried	White	162713.078888	80.000000
700	Other-service	State-gov	Unmarried	Black	149395.201919	15.000000

SQL for Real:
SELECT occupation,workclass,relationship,race,MIN(fnlwgt), COUNT(*) FROM C1 GROUP BY occupation,workclass,relationship,race

Resulted in 1160 records

SQL for Synthetic:
SELECT occupation,workclass,relationship,race,MIN(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY occupation,workclass,relationship,race

Resulted in 1122 records

Normalized Euclidean distance for (fnlwgt): 30.2

Hellinger Distance: 0.081

=====

Real				
	workclass	marital.status	occupation	MAX(capital) COUNT(*)
348	Self-emp-not-inc	Never-married	Adm-clerical	2174.000000 19.000000
319	Self-emp-not-inc	Divorced	Prof-specialty	99999.000000 69.000000
405	State-gov	Married-civ-spouse	Sales	99999.000000 9.000000
332	Self-emp-not-inc	Married-civ-spouse	Other-service	15024.000000 106.000000
350	Self-emp-not-inc	Never-married	Exec-managerial	99999.000000 48.000000

Synthetic				
	workclass	marital.status	occupation	MAX(capital) COUNT(*)
348	Self-emp-not-inc	Never-married	Adm-clerical	177.720661 19
319	Self-emp-not-inc	Divorced	Prof-specialty	283.114941 58
405	State-gov	Married-civ-spouse	Sales	58.701632 7
332	Self-emp-not-inc	Married-civ-spouse	Other-service	198.597981 108
350	Self-emp-not-inc	Never-married	Exec-managerial	193.928158 59

SQL for Real:
SELECT workclass,`marital.status`,occupation,MAX(capital), COUNT(*) FROM C1 GROUP BY workclass,`marital.status`,occupation

Resulted in 423 records

SQL for Synthetic:
SELECT workclass,`marital.status`,occupation,MAX(capital), COUNT(*) FROM C1_syn_06 GROUP BY workclass,`marital.status`,occupation

Resulted in 447 records

Normalized Euclidean distance for (capital): 19.52

Hellinger Distance: 0.049

=====

Real		
	income	MAX(age) COUNT(*)
0	<=50K	90 37155
1	>50K	90 11687

Synthetic		
	income	MAX(age) COUNT(*)
0	<=50K	90.000000 44080.000000
1	>50K	90.000000 4762.000000

SQL for Real:
SELECT income,MAX(age), COUNT(*) FROM C1 GROUP BY income

Resulted in 2 records

SQL for Synthetic:
SELECT income,MAX(age), COUNT(*) FROM C1_syn_06 GROUP BY income

Resulted in 2 records

Normalized Euclidean distance for (age): 1.41

Hellinger Distance: 0.137

=====

Real							
	sex	race	education	income	native.country	relationship	MAX(age) COUNT(*)
1505	Female	White	Masters	<=50K	Peru	Unmarried	nan nan
2951	Male	White	Assoc-voc	<=50K	Peru	Own-child	nan nan
1089	Female	White	Assoc-voc	<=50K	Ireland	Wife	nan nan
1835	Male	Asian-Pac-Islander	5th-6th	<=50K	United-States	Husband	78.000000 2.000000
1628	Female	White	Some-college	<=50K	Holand-Netherlands	Unmarried	nan nan

Synthetic							
	sex	race	education	income	native.country	relationship	MAX(age) COUNT(*)
1505	Female	White	Masters	<=50K	Peru	Unmarried	47 4
2951	Male	White	Assoc-voc	<=50K	Peru	Own-child	21 1
1089	Female	White	Assoc-voc	<=50K	Ireland	Wife	30 1
1835	Male	Asian-Pac-Islander	5th-6th	<=50K	United-States	Husband	71 2
1628	Female	White	Some-college	<=50K	Holand-Netherlands	Unmarried	24 3

SQL for Real:
SELECT sex,race,education,income,`native.country`,relationship,MAX(age), COUNT(*) FROM C1 GROUP BY sex,race,education,income,`native.country`,relationship

Resulted in 2920 records

SQL for Synthetic:
SELECT sex,race,education,income,`native.country`,relationship,MAX(age), COUNT(*) FROM C1_syn_06 GROUP BY sex,race,education,income,`native.country`,relationship

Resulted in 3549 records

Normalized Euclidean distance for (age): 34.89

Hellinger Distance: 0.361

=====

Real								
	race	income	workclass	education	marital.status	occupation	native.country	SUM(`hours.per.week`) COUNT(*)
6878	White	<=50K	Private	HS-grad	Divorced	Adm-clerical	Ecuador	nan nan
7201	White	<=50K	Private	HS-grad	Married-civ-spouse	Transport-moving	Holand-Netherlands	nan nan
11469	White	>50K	Self-emp-not-inc	HS-grad	Married-civ-spouse	Farming-fishing	United-States	2087.000000 36.000000
3347	White	<=50K	?	HS-grad	Married-civ-spouse	?	Columbia	nan nan
7123	White	<=50K	Private	HS-grad	Married-civ-spouse	Machine-op-inspct	Puerto-Rico	90.000000 2.000000

Synthetic								
	race	income	workclass	education	marital.status	occupation	native.country	SUM(`hours.per.week`) COUNT(*)
6878	White	<=50K	Private	HS-grad	Divorced	Adm-clerical	Ecuador	40.025527 1
7201	White	<=50K	Private	HS-grad	Married-civ-spouse	Transport-moving	Holand-Netherlands	40.031101 1
11469	White	>50K	Self-emp-not-inc	HS-grad	Married-civ-spouse	Farming-fishing	United-States	322.841140 8
3347	White	<=50K	?	HS-grad	Married-civ-spouse	?	Columbia	280.117099 7
7123	White	<=50K	Private	HS-grad	Married-civ-spouse	Machine-op-inspct	Puerto-Rico	5639.492177 141

SQL for Real:
SELECT race,income,workclass,education,`marital.status`,occupation,`native.country`,SUM(`hours.per.week`), COUNT(*) FROM C1 GROUP BY race,income,workclass,education,`marital.status`,occupation,`native.country`
Resulted in 8403 records

SQL for Synthetic:
SELECT race,income,workclass,education,`marital.status`,occupation,`native.country`,SUM(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY race,income,workclass,education,`marital.status`,occupation,`native.country`
Resulted in 11619 records

Normalized Euclidean distance for (`hours.per.week`): 55.59

Hellinger Distance: 0.265

=====

Real			
	education	AVG(`hours.per.week`)	COUNT(*)
2	12th	35.374429	657
5	7th-8th	39.003141	955
11	HS-grad	40.640775	15784
1	11th	33.952539	1812
8	Assoc-voc	41.658418	2061

Synthetic			
	education	AVG(`hours.per.week`)	COUNT(*)
2	12th	39.758112	626.000000
5	7th-8th	39.877692	781.000000
11	HS-grad	39.993373	16455.000000
1	11th	39.755015	1694.000000
8	Assoc-voc	40.023089	1935.000000

SQL for Real:
SELECT education,AVG(`hours.per.week`), COUNT(*) FROM C1 GROUP BY education
Resulted in 16 records

SQL for Synthetic:
SELECT education,AVG(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY education
Resulted in 16 records

Normalized Euclidean distance for (`hours.per.week`): 4.0

Hellinger Distance: 0.026

=====

Real			
	occupation	AVG(`hours.per.week`)	COUNT(*)
13	Tech-support	39.720609	1446
3	Craft-repair	42.269306	6112
1	Adm-clerical	37.708073	5611
6	Handlers-cleaners	37.911197	2072
10	Prof-specialty	42.282890	6172

Synthetic			
	occupation	AVG(`hours.per.week`)	COUNT(*)
13	Tech-support	39.971780	1463.000000
3	Craft-repair	40.002931	5926.000000
1	Adm-clerical	39.902178	5729.000000
6	Handlers-cleaners	39.878155	2129.000000
10	Prof-specialty	40.051494	6213.000000

SQL for Real:
SELECT occupation,AVG(`hours.per.week`), COUNT(*) FROM C1 GROUP BY occupation
Resulted in 15 records

SQL for Synthetic:
SELECT occupation,AVG(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY occupation
Resulted in 15 records

Normalized Euclidean distance for (`hours.per.week`): 3.87

Hellinger Distance: 0.027

=====

Real							
	occupation	sex	native.country	education	relationship	AVG(capital)	COUNT(*)
4646	Prof-specialty	Female	Germany	Assoc-acdm	Own-child	nan	nan
2606	Farming-fishing	Male	Laos	10th	Own-child	nan	nan
839	Adm-clerical	Female	Puerto-Rico	Masters	Wife	nan	nan
4224	Other-service	Male	Ireland	Bachelors	Husband	nan	nan
2993	Handlers-cleaners	Male	Outlying-US(Guam-USVI-etc)	HS-grad	Husband	nan	nan

Synthetic							
	occupation	sex	native.country	education	relationship	AVG(capital)	COUNT(*)
4646	Prof-specialty	Female	Germany	Assoc-acdm	Own-child	25.533381	1
2606	Farming-fishing	Male	Laos	10th	Own-child	-92.751280	1
839	Adm-clerical	Female	Puerto-Rico	Masters	Wife	-13.981253	5
4224	Other-service	Male	Ireland	Bachelors	Husband	62.192245	1
2993	Handlers-cleaners	Male	Outlying-US(Guam-USVI-etc)	HS-grad	Husband	139.769567	1

SQL for Real:
SELECT occupation,sex,`native.country`,education,relationship,AVG(capital), COUNT(*) FROM C1 GROUP BY occupation,sex,`native.country`,education,relationship
Resulted in 4565 records

SQL for Synthetic:
SELECT occupation,sex,`native.country`,education,relationship,AVG(capital), COUNT(*) FROM C1_syn_06 GROUP BY occupation,sex,`native.country`,education,relationship
Resulted in 6788 records

Normalized Euclidean distance for (capital): 44.73

Hellinger Distance: 0.268

=====

Real										
	race	relationship	income	sex	native.country	marital.status	occupation	education	MIN(`hours.per.week`)	COUNT(*)
10220	White	Own-child	<=50K	Male	Peru	Never-married	Handlers-cleaners	11th	20.000000	1.000000
2783	Black	Unmarried	<=50K	Female	United-States	Divorced	Tech-support	HS-grad	35.000000	2.000000
8118	White	Not-in-family	<=50K	Male	United-States	Never-married	Machine-op-inspct	HS-grad	10.000000	93.000000
9650	White	Own-child	<=50K	Female	Thailand	Married-civ-spouse	?	11th	nan	nan
12250	White	Unmarried	<=50K	Male	United-States	Never-married	Machine-op-inspct	10th	40.000000	1.000000

Synthetic										
	race	relationship	income	sex	native.country	marital.status	occupation	education	MIN(`hours.per.week`)	COUNT(*)
10220	White	Own-child	<=50K	Male	Peru	Never-married	Handlers-cleaners	11th	39.526037	2
2783	Black	Unmarried	<=50K	Female	United-States	Divorced	Tech-support	HS-grad	39.956414	5
8118	White	Not-in-family	<=50K	Male	United-States	Never-married	Machine-op-inspct	HS-grad	39.897268	71
9650	White	Own-child	<=50K	Female	Thailand	Married-civ-spouse	?	11th	39.497325	1
12250	White	Unmarried	<=50K	Male	United-States	Never-married	Machine-op-inspct	10th	40.034386	1

SQL for Real:
SELECT race,relationship,income,sex,`native.country`,`marital.status`,occupation,education,MIN(`hours.per.week`), COUNT(*) FROM C1 GROUP BY race,relationship,income,sex,`native.country`,`marital.status`,occupation,education
Resulted in 9390 records

SQL for Synthetic:
SELECT race,relationship,income,sex,`native.country`,`marital.status`,occupation,education,MIN(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY race,relationship,income,sex,`native.country`,`marital.status`,occupation,education
Resulted in 13048 records

Normalized Euclidean distance for (`hours.per.week`): 57.52

Hellinger Distance: 0.264

=====

Real						
	sex	marital.status	race	education	income	MIN(`hours.per.week`) COUNT(*)
445	Female	Widowed	White	Assoc-acdm	<=50K	3 27
119	Female	Married-civ-spouse	Asian-Pac-Islander	7th-8th	<=50K	48 2
126	Female	Married-civ-spouse	Asian-Pac-Islander	Bachelors	<=50K	5 20
12	Female	Divorced	Asian-Pac-Islander	11th	<=50K	15 1
521	Male	Divorced	White	10th	<=50K	12 56

Synthetic						
	sex	marital.status	race	education	income	MIN(`hours.per.week`) COUNT(*)
445	Female	Widowed	White	Assoc-acdm	<=50K	38.958373 21.000000
119	Female	Married-civ-spouse	Asian-Pac-Islander	7th-8th	<=50K	40.010611 3.000000
126	Female	Married-civ-spouse	Asian-Pac-Islander	Bachelors	<=50K	39.530760 24.000000
12	Female	Divorced	Asian-Pac-Islander	11th	<=50K	nan nan
521	Male	Divorced	White	10th	<=50K	39.005802 69.000000

SQL for Real:
SELECT sex,`marital.status`,race,education,income,MIN(`hours.per.week`), COUNT(*) FROM C1 GROUP BY sex,`marital.status`,race,education,income
Resulted in 948 records

SQL for Synthetic:
SELECT sex,`marital.status`,race,education,income,MIN(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY sex,`marital.status`,race,education,income
Resulted in 774 records

Normalized Euclidean distance for (`hours.per.week`): 26.27

Hellinger Distance: 0.173

=====

Real					
	sex	workclass	relationship	MIN(`hours.per.week`)	COUNT(*)
8	Female	Federal-gov	Own-child	10.000000	42.000000
14	Female	Local-gov	Own-child	2.000000	170.000000
66	Male	Self-emp-inc	Other-relative	25.000000	10.000000
63	Male	Private	Wife	40.000000	2.000000
26	Female	Self-emp-inc	Own-child	8.000000	25.000000

Synthetic					
	sex	workclass	relationship	MIN(`hours.per.week`)	COUNT(*)
8	Female	Federal-gov	Own-child	38.944736	60
14	Female	Local-gov	Own-child	38.925084	177
66	Male	Self-emp-inc	Other-relative	39.962764	10
63	Male	Private	Wife	39.969150	6
26	Female	Self-emp-inc	Own-child	38.974765	24

SQL for Real:
SELECT sex,workclass,relationship,MIN(`hours.per.week`), COUNT(*) FROM C1 GROUP BY sex,workclass,relationship
Resulted in 73 records

SQL for Synthetic:
SELECT sex,workclass,relationship,MIN(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY sex,workclass,relationship
Resulted in 80 records

Normalized Euclidean distance for (`hours.per.week`): 8.54

Hellinger Distance: 0.029

=====

Real										
	relationship	race	occupation	education	native.country	sex	income	marital.status	MAX(capital)	COUNT(*)
12567	Wife	White	Adm-clerical	Some-college	Peru	Female	<=50K	Married-civ-spouse	nan	nan
3614	Not-in-family	White	?	Assoc-acdm	Peru	Female	<=50K	Never-married	nan	nan
2985	Not-in-family	Black	Craft-repair	10th	Puerto-Rico	Male	<=50K	Divorced	nan	nan
4645	Not-in-family	White	Exec-managerial	Masters	China	Female	<=50K	Never-married	nan	nan
6986	Other-relative	White	?	7th-8th	Puerto-Rico	Female	<=50K	Never-married	nan	nan

Synthetic										
	relationship	race	occupation	education	native.country	sex	income	marital.status	MAX(capital)	COUNT(*)
12567	Wife	White	Adm-clerical	Some-college	Peru	Female	<=50K	Married-civ-spouse	-58.418351	1
3614	Not-in-family	White	?	Assoc-acdm	Peru	Female	<=50K	Never-married	12.795618	1
2985	Not-in-family	Black	Craft-repair	10th	Puerto-Rico	Male	<=50K	Divorced	52.081197	1
4645	Not-in-family	White	Exec-managerial	Masters	China	Female	<=50K	Never-married	-67.791015	1
6986	Other-relative	White	?	7th-8th	Puerto-Rico	Female	<=50K	Never-married	148.859027	1

SQL for Real:
SELECT relationship,race,occupation,education,`native.country`,sex,income,`marital.status`,MAX(capital), COUNT(*) FROM C1 GROUP BY relationship,race,occupation,education,`native.country`,sex,income,`marital.status`
Resulted in 9390 records

SQL for Synthetic:
SELECT relationship,race,occupation,education,`native.country`,sex,income,`marital.status`,MAX(capital), COUNT(*) FROM C1_syn_06 GROUP BY relationship,race,occupation,education,`native.country`,sex,income,`marital.status`
Resulted in 13048 records

Normalized Euclidean distance for (capital): 57.52

Hellinger Distance: 0.264

=====

Real									
	race	relationship	marital.status	income	workclass	native.country	education	MAX(fnlwgt)	COUNT(*)
3360	White	Husband	Married-civ-spouse	>50K	Private	Germany	Bachelors	389843.000000	7.000000
5155	White	Other-relative	Divorced	<=50K	Local-gov	United-States	HS-grad	323829.000000	3.000000
3148	White	Husband	Married-civ-spouse	<=50K	Self-emp-not-inc	United-States	10th	439777.000000	43.000000
3541	White	Husband	Married-civ-spouse	>50K	State-gov	Hungary	Bachelors	nan	nan
1380	Black	Not-in-family	Widowed	<=50K	State-gov	United-States	Masters	nan	nan

		Synthetic							
	race	relationship	marital.status	income	workclass	native.country	education	MAX(fnlwgt)	COUNT(*)
3360	White	Husband	Married-civ-spouse	>50K	Private	Germany	Bachelors	184139.440113	1
5155	White	Other-relative	Divorced	<=50K	Local-gov	United-States	HS-grad	173653.212975	1
3148	White	Husband	Married-civ-spouse	<=50K	Self-emp-not-inc	United-States	10th	190573.127290	38
3541	White	Husband	Married-civ-spouse	>50K	State-gov	Hungary	Bachelors	180833.461886	1
1380	Black	Not-in-family	Widowed	<=50K	State-gov	United-States	Masters	161701.469662	1

SQL for Real:

SELECT race,relationship,`marital.status`,income,workclass,`native.country`,education,MAX(fnlwgt), COUNT(*) FROM C1 GROUP BY race,relationship,`marital.status`,income,workclass,`native.country`,education

Resulted in 5799 records

SQL for Synthetic:

SELECT race,relationship,`marital.status`,income,workclass,`native.country`,education,MAX(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY race,relationship,`marital.status`,income,workclass,`native.country`,education

Resulted in 7748 records

Normalized Euclidean distance for (fnlwgt): 46.68

Hellinger Distance: 0.319

=====

		Real	
	workclass	SUM(fnlwgt)	COUNT(*)
6	State-gov	360410194	1981
2	Local-gov	596345319	3136
4	Self-emp-inc	303388390	1695
0	?	529802400	2830
1	Federal-gov	262900921	1432

		Synthetic	
	workclass	SUM(fnlwgt)	COUNT(*)
6	State-gov	338630858.672139	1982.000000
2	Local-gov	567199972.519647	3105.000000
4	Self-emp-inc	284019318.551430	1645.000000
0	?	506616985.296546	2844.000000
1	Federal-gov	249839978.066589	1426.000000

SQL for Real:

SELECT workclass,SUM(fnlwgt), COUNT(*) FROM C1 GROUP BY workclass

Resulted in 7 records

SQL for Synthetic:

SELECT workclass,SUM(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY workclass

Resulted in 7 records

Normalized Euclidean distance for (fnlwgt): 2.65

Hellinger Distance: 0.003

=====

		Real				
	education	marital.status	income	relationship	MAX(capital)	COUNT(*)
181	9th	Divorced	>50K	Unmarried	99999	1
223	Assoc-acdm	Married-spouse-absent	<=50K	Unmarried	0	8
53	11th	Never-married	<=50K	Not-in-family	4416	170
56	11th	Never-married	<=50K	Unmarried	6849	73
416	Masters	Divorced	>50K	Unmarried	99999	40

		Synthetic				
	education	marital.status	income	relationship	MAX(capital)	COUNT(*)
181	9th	Divorced	>50K	Unmarried	-77.801174	1.000000
223	Assoc-acdm	Married-spouse-absent	<=50K	Unmarried	168.139639	5.000000
53	11th	Never-married	<=50K	Not-in-family	266.105654	130.000000
56	11th	Never-married	<=50K	Unmarried	181.215581	74.000000
416	Masters	Divorced	>50K	Unmarried	84.521630	11.000000

SQL for Real:

SELECT education,`marital.status`,income,relationship,MAX(capital), COUNT(*) FROM C1 GROUP BY education,`marital.status`,income,relationship

Resulted in 547 records

SQL for Synthetic:

SELECT education,`marital.status`,income,relationship,MAX(capital), COUNT(*) FROM C1_syn_06 GROUP BY education,`marital.status`,income,relationship

Resulted in 535 records

Normalized Euclidean distance for (capital): 20.9

Hellinger Distance: 0.166

=====

Real						
	marital.status	sex	native.country	race	SUM(`hours.per.week`)	COUNT(*)
34	Divorced	Female	Jamaica	Black	276.000000	7.000000
481	Never-married	Female	United-States	White	184395.000000	5361.000000
650	Separated	Male	Hungary	Black	nan	nan
260	Married-civ-spouse	Male	Ireland	Asian-Pac-Islander	nan	nan
249	Married-civ-spouse	Male	Holand-Netherlands	Black	nan	nan

Synthetic						
	marital.status	sex	native.country	race	SUM(`hours.per.week`)	COUNT(*)
34	Divorced	Female	Jamaica	Black	160.040880	4
481	Never-married	Female	United-States	White	172040.547291	4324
650	Separated	Male	Hungary	Black	40.000658	1
260	Married-civ-spouse	Male	Ireland	Asian-Pac-Islander	120.063580	3
249	Married-civ-spouse	Male	Holand-Netherlands	Black	80.000798	2

SQL for Real:
SELECT `marital.status`,`sex`,`native.country`,`race`,SUM(`hours.per.week`), COUNT(*) FROM C1 GROUP BY `marital.status`,`sex`,`native.country`,`race`
Resulted in 658 records

SQL for Synthetic:
SELECT `marital.status`,`sex`,`native.country`,`race`,SUM(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY `marital.status`,`sex`,`native.country`,`race`
Resulted in 739 records

Normalized Euclidean distance for (`hours.per.week`): 19.44

Hellinger Distance: 0.395

=====

Real											
	native.country	income	race	marital.status	occupation	relationship	workclass	education	sex	MIN(age)	COUNT(*)
8332	Puerto-Rico	<=50K	White	Married-spouse-absent	Sales	Not-in-family	Private	Masters	Male	nan	nan
14474	United-States	<=50K	White	Never-married	Craft-repair	Not-in-family	Private	Masters	Female	39.000000	1.000000
3334	India	<=50K	White	Never-married	Sales	Other-relative	Private	Some-college	Female	nan	nan
6246	Puerto-Rico	<=50K	Amer-Indian-Eskimo	Divorced	Prof-specialty	Not-in-family	Private	Masters	Female	nan	nan
9575	Puerto-Rico	>50K	White	Married-civ-spouse	?	Husband	?	Assoc-voc	Male	nan	nan

Synthetic											
	native.country	income	race	marital.status	occupation	relationship	workclass	education	sex	MIN(age)	COUNT(*)
8332	Puerto-Rico	<=50K	White	Married-spouse-absent	Sales	Not-in-family	Private	Masters	Male	49	2
14474	United-States	<=50K	White	Never-married	Craft-repair	Not-in-family	Private	Masters	Female	35	1
3334	India	<=50K	White	Never-married	Sales	Other-relative	Private	Some-college	Female	28	1
6246	Puerto-Rico	<=50K	Amer-Indian-Eskimo	Divorced	Prof-specialty	Not-in-family	Private	Masters	Female	58	1
9575	Puerto-Rico	>50K	White	Married-civ-spouse	?	Husband	?	Assoc-voc	Male	66	1

SQL for Real:
SELECT `native.country`,`income`,`race`,`marital.status`,`occupation`,`relationship`,`workclass`,`education`,`sex`,MIN(age), COUNT(*) FROM C1 GROUP BY `native.country`,`income`,`race`,`marital.status`,`occupation`,`relationship`,`workclass`,`education`,`sex`
Resulted in 12664 records

SQL for Synthetic:
SELECT `native.country`,`income`,`race`,`marital.status`,`occupation`,`relationship`,`workclass`,`education`,`sex`,MIN(age), COUNT(*) FROM C1_syn_06 GROUP BY `native.country`,`income`,`race`,`marital.status`,`occupation`,`relationship`,`workclass`,`education`,`sex`
Resulted in 17127 records

Normalized Euclidean distance for (age): 64.61

Hellinger Distance: 0.252

=====

Real					
	occupation	relationship	income	sex	MIN(`hours.per.week`)
100	Handlers-cleaners	Husband	<=50K	Male	2
84	Farming-fishing	Husband	>50K	Male	2
195	Protective-serv	Own-child	<=50K	Male	3
85	Farming-fishing	Not-in-family	<=50K	Female	14
162	Priv-house-serv	Unmarried	<=50K	Female	4

Synthetic					
	occupation	relationship	income	sex	MIN(`hours.per.week`)
100	Handlers-cleaners	Husband	<=50K	Male	38.949391
84	Farming-fishing	Husband	>50K	Male	39.163313
195	Protective-serv	Own-child	<=50K	Male	39.478496
85	Farming-fishing	Not-in-family	<=50K	Female	39.941380
162	Priv-house-serv	Unmarried	<=50K	Female	39.044383

SQL for Real:
SELECT occupation,relationship,income,sex,MIN(`hours.per.week`), COUNT(*) FROM C1 GROUP BY occupation,relationship,income,sex
Resulted in 261 records

SQL for Synthetic:
SELECT occupation,relationship,income,sex,MIN(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY occupation,relationship,income,sex
Resulted in 258 records

Normalized Euclidean distance for (`hours.per.week`): 15.2

Hellinger Distance: 0.161

=====

Real										
	race	income	education	sex	relationship	workclass	occupation	marital.status	native.country	AVG(fnlwgt) COUNT(*)
16929	White	>50K	Some-college	Female	Wife	Private	Machine-op-inspct	Married-civ-spouse	United-States	278568.500000 4.000000
3101	Black	<=50K	Some-college	Female	Not-in-family	State-gov	Adm-clerical	Married-spouse-absent	United-States	nan nan
15987	White	>50K	Bachelors	Female	Wife	Federal-gov	Exec-managerial	Married-civ-spouse	United-States	213226.000000 2.000000
6224	White	<=50K	9th	Male	Husband	Private	Craft-repair	Married-civ-spouse	Peru	nan nan
9589	White	<=50K	Doctorate	Male	Unmarried	State-gov	Exec-managerial	Never-married	United-States	nan nan

Synthetic										
	race	income	education	sex	relationship	workclass	occupation	marital.status	native.country	AVG(fnlwgt) COUNT(*)
16929	White	>50K	Some-college	Female	Wife	Private	Machine-op-inspct	Married-civ-spouse	United-States	175637.916397 3
3101	Black	<=50K	Some-college	Female	Not-in-family	State-gov	Adm-clerical	Married-spouse-absent	United-States	158523.784868 1
15987	White	>50K	Bachelors	Female	Wife	Federal-gov	Exec-managerial	Married-civ-spouse	United-States	144973.920821 1
6224	White	<=50K	9th	Male	Husband	Private	Craft-repair	Married-civ-spouse	Peru	179030.389533 3
9589	White	<=50K	Doctorate	Male	Unmarried	State-gov	Exec-managerial	Never-married	United-States	147852.275859 1

SQL for Real:
SELECT race,income,education,sex,relationship,workclass,occupation,`marital.status`,`native.country`,AVG(fnlwgt), COUNT(*) FROM C1 GROUP BY race,income,education,sex,relationship,workclass,occupation,`marital.status`,`native.country`
Resulted in 12664 records

SQL for Synthetic:
SELECT race,income,education,sex,relationship,workclass,occupation,`marital.status`,`native.country`,AVG(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY race,income,education,sex,relationship,workclass,occupation,`marital.status`,`native.country`
Resulted in 17127 records

Normalized Euclidean distance for (fnlwgt): 64.61

Hellinger Distance: 0.252

=====

Real				
	income	workclass	MIN(`hours.per.week`)	COUNT(*)
7	>50K	?	1	267
2	<=50K	Local-gov	2	2209
10	>50K	Private	1	7387
3	<=50K	Private	1	26519
0	<=50K	?	1	2563

Synthetic				
	income	workclass	MIN(`hours.per.week`)	COUNT(*)
7	>50K	?	38.897276	95.000000
2	<=50K	Local-gov	38.913042	2683.000000
10	>50K	Private	38.952368	2930.000000
3	<=50K	Private	38.883397	31092.000000
0	<=50K	?	38.865705	2749.000000

SQL for Real:
SELECT income,workclass,MIN(`hours.per.week`), COUNT(*) FROM C1 GROUP BY income,workclass
Resulted in 14 records

SQL for Synthetic:
SELECT income,workclass,MIN(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY income,workclass
Resulted in 14 records

Normalized Euclidean distance for (`hours.per.week`): 3.74

Hellinger Distance: 0.138

=====

Real								
	workclass	education	sex	race	native.country	income	MAX(`hours.per.week`)	COUNT(*)
2305	Private	HS-grad	Male	White	Hungary	>50K	nan	nan
634	Federal-gov	Masters	Male	White	Peru	<=50K	nan	nan
2281	Private	HS-grad	Male	Other	Ireland	<=50K	nan	nan
1166	Local-gov	Some-college	Female	Black	United-States	<=50K	55.000000	46.000000
250	?	HS-grad	Female	Black	?	<=50K	nan	nan

Synthetic								
	workclass	education	sex	race	native.country	income	MAX(`hours.per.week`)	COUNT(*)
2305	Private	HS-grad	Male	White	Hungary	>50K	40.430558	23
634	Federal-gov	Masters	Male	White	Peru	<=50K	40.003294	5
2281	Private	HS-grad	Male	Other	Ireland	<=50K	39.986253	1
1166	Local-gov	Some-college	Female	Black	United-States	<=50K	40.047980	24
250	?	HS-grad	Female	Black	?	<=50K	40.014614	3

SQL for Real:
SELECT workclass,education,sex,race,`native.country`,income,MAX(`hours.per.week`), COUNT(*) FROM C1 GROUP BY workclass,education,sex,race,`native.country`,income
Resulted in 2913 records

SQL for Synthetic:
SELECT workclass,education,sex,race,`native.country`,income,MAX(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY workclass,education,sex,race,`native.country`,income
Resulted in 3665 records

Normalized Euclidean distance for (`hours.per.week`): 35.11

Hellinger Distance: 0.358

=====

Real			
	workclass	MAX(age)	COUNT(*)
4	Self-emp-inc	85	1695
3	Private	90	33906
2	Local-gov	90	3136
6	State-gov	81	1981
5	Self-emp-not-inc	90	3862

Synthetic			
	workclass	MAX(age)	COUNT(*)
4	Self-emp-inc	84.000000	1645.000000
3	Private	90.000000	34022.000000
2	Local-gov	90.000000	3105.000000
6	State-gov	81.000000	1982.000000
5	Self-emp-not-inc	90.000000	3818.000000

SQL for Real:
SELECT workclass,MAX(age), COUNT(*) FROM C1 GROUP BY workclass
Resulted in 7 records

SQL for Synthetic:
SELECT workclass,MAX(age), COUNT(*) FROM C1_syn_06 GROUP BY workclass
Resulted in 7 records

Normalized Euclidean distance for (age): 2.65

Hellinger Distance: 0.003

=====

Real					
	relationship	race	education	SUM(age)	COUNT(*)
382	Wife	Other	7th-8th	81	2
158	Other-relative	Asian-Pac-Islander	11th	69	2
167	Other-relative	Asian-Pac-Islander	Prof-school	90	2
45	Husband	Other	11th	206	6
139	Not-in-family	White	12th	3473	96

Synthetic					
	relationship	race	education	SUM(age)	COUNT(*)
382	Wife	Other	7th-8th	nan	nan
158	Other-relative	Asian-Pac-Islander	11th	53.000000	3.000000
167	Other-relative	Asian-Pac-Islander	Prof-school	28.000000	1.000000
45	Husband	Other	11th	69.000000	2.000000
139	Not-in-family	White	12th	3634.000000	101.000000

SQL for Real:
SELECT relationship,race,education,SUM(age), COUNT(*) FROM C1 GROUP BY relationship,race,education
Resulted in 404 records

SQL for Synthetic:
SELECT relationship,race,education,SUM(age), COUNT(*) FROM C1_syn_06 GROUP BY relationship,race,education
Resulted in 371 records

Normalized Euclidean distance for (age): 18.71

Hellinger Distance: 0.064

=====

Real								
	sex	occupation	race	marital.status	native.country	education	AVG(fnlwgt)	COUNT(*)
170	Female	?	White	Divorced	Scotland	Assoc-voc	nan	nan
2777	Female	Priv-house-serv	Other	Divorced	Guatemala	12th	nan	nan
8101	Male	Sales	Other	Never-married	United-States	Masters	nan	nan
8319	Male	Sales	White	Never-married	Holand-Netherlands	HS-grad	nan	nan
8348	Male	Sales	White	Never-married	Outlying-US(Guam-USVI-etc)	HS-grad	nan	nan

Synthetic								
	sex	occupation	race	marital.status	native.country	education	AVG(fnlwgt)	COUNT(*)
170	Female	?	White	Divorced	Scotland	Assoc-voc	180578.478483	1
2777	Female	Priv-house-serv	Other	Divorced	Guatemala	12th	184754.137780	1
8101	Male	Sales	Other	Never-married	United-States	Masters	180284.705271	1
8319	Male	Sales	White	Never-married	Holand-Netherlands	HS-grad	201563.361192	2
8348	Male	Sales	White	Never-married	Outlying-US(Guam-USVI-etc)	HS-grad	200641.749710	2

SQL for Real:

SELECT sex,occupation,race,`marital.status`,`native.country`,education,AVG(fnlwgt), COUNT(*) FROM C1 GROUP BY sex,occupation,race,`marital.status`,`native.country`,education

Resulted in 6163 records

SQL for Synthetic:

SELECT sex,occupation,race,`marital.status`,`native.country`,education,AVG(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY sex,occupation,race,`marital.status`,`native.country`,education

Resulted in 8996 records

Normalized Euclidean distance for (fnlwgt): 49.57

Hellinger Distance: 0.252

=====

Real									
	marital.status	education	race	income	workclass	relationship	sex	native.country	SUM(`hours.per.week`) COUNT(*)
5737	Never-married	Assoc-voc	White	<=50K	Private	Other-relative	Female	United-States	72.000000 2.000000
165	Divorced	11th	White	<=50K	Self-emp-not-inc	Unmarried	Female	Puerto-Rico	nan nan
1045	Divorced	HS-grad	White	<=50K	Federal-gov	Husband	Male	Puerto-Rico	nan nan
6359	Never-married	HS-grad	Black	<=50K	?	Own-child	Female	Ireland	nan nan
2093	Married-civ-spouse	1st-4th	Black	<=50K	?	Husband	Male	United-States	40.000000 1.000000

Synthetic									
	marital.status	education	race	income	workclass	relationship	sex	native.country	SUM(`hours.per.week`) COUNT(*)
5737	Never-married	Assoc-voc	White	<=50K	Private	Other-relative	Female	United-States	119.410136 3
165	Divorced	11th	White	<=50K	Self-emp-not-inc	Unmarried	Female	Puerto-Rico	39.967635 1
1045	Divorced	HS-grad	White	<=50K	Federal-gov	Husband	Male	Puerto-Rico	40.091665 1
6359	Never-married	HS-grad	Black	<=50K	?	Own-child	Female	Ireland	40.001215 1
2093	Married-civ-spouse	1st-4th	Black	<=50K	?	Husband	Male	United-States	38.960631 1

SQL for Real:

SELECT `marital.status`,education,race,income,workclass,relationship,sex,`native.country`,SUM(`hours.per.week`), COUNT(*) FROM C1 GROUP BY `marital.status`,education,race,income,workclass,relationship,sex,`native.country`

Resulted in 6863 records

SQL for Synthetic:

SELECT `marital.status`,education,race,income,workclass,relationship,sex,`native.country`,SUM(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY `marital.status`,education,race,income,workclass,relationship,sex,`native.country`

Resulted in 9165 records

Normalized Euclidean distance for (`hours.per.week`): 50.11

Hellinger Distance: 0.311

=====

Real							
	sex	education	native.country	occupation	race	AVG(age)	COUNT(*)
4318	Male	Masters	Germany	Craft-repair	White	nan	nan
811	Female	Assoc-voc	Philippines	Machine-op-inspct	White	nan	nan
2753	Male	5th-6th	Peru	Craft-repair	Black	nan	nan
4900	Male	Some-college	United-States	Armed-Forces	White	25.666667	3.000000
1803	Female	Prof-school	Cuba	Prof-specialty	White	nan	nan

Synthetic							
	sex	education	native.country	occupation	race	AVG(age)	COUNT(*)
4318	Male	Masters	Germany	Craft-repair	White	43.000000	1
811	Female	Assoc-voc	Philippines	Machine-op-inspct	White	47.000000	1
2753	Male	5th-6th	Peru	Craft-repair	Black	57.000000	1
4900	Male	Some-college	United-States	Armed-Forces	White	30.750000	4
1803	Female	Prof-school	Cuba	Prof-specialty	White	67.000000	1

SQL for Real:

SELECT sex,education,`native.country`,occupation,race,AVG(age), COUNT(*) FROM C1 GROUP BY sex,education,`native.country`,occupation,race

Resulted in 3582 records

SQL for Synthetic:

SELECT sex,education,`native.country`,occupation,race,AVG(age), COUNT(*) FROM C1_syn_06 GROUP BY sex,education,`native.country`,occupation,race

Resulted in 4954 records

Normalized Euclidean distance for (age): 39.13

Hellinger Distance: 0.306

=====

Real						
	workclass	occupation	relationship	native.country	SUM(`hours.per.week`)	COUNT(*)
2268	Self-emp-inc	Exec-managerial	Husband	Peru	nan	nan
2990	State-gov	Protective-serv	Not-in-family	?	nan	nan
1449	Private	Machine-op-inspct	Not-in-family	Columbia	120.000000	3.000000
1307	Private	Farming-fishing	Unmarried	China	nan	nan
1631	Private	Other-service	Own-child	Germany	90.000000	4.000000

Synthetic

	workclass	occupation	relationship	native.country	SUM(`hours.per.week`)	COUNT(*)
2268	Self-emp-inc	Exec-managerial	Husband	Peru	241.988902	6
2990	State-gov	Protective-serv	Not-in-family	?	119.995860	3
1449	Private	Machine-op-inspct	Not-in-family	Columbia	759.862730	19
1307	Private	Farming-fishing	Unmarried	China	40.046355	1
1631	Private	Other-service	Own-child	Germany	40.015598	1

SQL for Real:
SELECT workclass,occupation,relationship,`native.country`,SUM(`hours.per.week`), COUNT(*) FROM C1 GROUP BY workclass,occupation,relationship,`native.country`
Resulted in 2330 records

SQL for Synthetic:
SELECT workclass,occupation,relationship,`native.country`,SUM(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY workclass,occupation,relationship,`native.country`
Resulted in 3049 records

Normalized Euclidean distance for (`hours.per.week`): 34.58
Hellinger Distance: 0.337

=====

Real											
	native.country	income	race	marital.status	sex	workclass	occupation	relationship	education	MIN(fnlwgt)	COUNT(*)
696	?	<=50K	White	Never-married	Male	Private	Protective-serv	Own-child	HS-grad	nan	nan
1614	Columbia	<=50K	White	Married-civ-spouse	Male	State-gov	Farming-fishing	Husband	HS-grad	nan	nan
4344	Japan	<=50K	White	Never-married	Male	Private	Adm-clerical	Not-in-family	HS-grad	nan	nan
10148	Scotland	<=50K	White	Never-married	Female	Private	Tech-support	Not-in-family	Bachelors	nan	nan
8541	Puerto-Rico	<=50K	White	Never-married	Female	Private	Protective-serv	Own-child	10th	nan	nan

Synthetic											
	native.country	income	race	marital.status	sex	workclass	occupation	relationship	education	MIN(fnlwgt)	COUNT(*)
696	?	<=50K	White	Never-married	Male	Private	Protective-serv	Own-child	HS-grad	197984.034859	1
1614	Columbia	<=50K	White	Married-civ-spouse	Male	State-gov	Farming-fishing	Husband	HS-grad	177869.045179	1
4344	Japan	<=50K	White	Never-married	Male	Private	Adm-clerical	Not-in-family	HS-grad	178049.212330	1
10148	Scotland	<=50K	White	Never-married	Female	Private	Tech-support	Not-in-family	Bachelors	179754.398381	1
8541	Puerto-Rico	<=50K	White	Never-married	Female	Private	Protective-serv	Own-child	10th	168957.692590	1

SQL for Real:
SELECT `native.country`,income,race,`marital.status`,sex,workclass,occupation,relationship,education,MIN(fnlwgt), COUNT(*) FROM C1 GROUP BY `native.country`,income,race,`marital.status`,sex,workclass,occupation,relationship,education
Resulted in 12664 records

SQL for Synthetic:
SELECT `native.country`,income,race,`marital.status`,sex,workclass,occupation,relationship,education,MIN(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY `native.country`,income,race,`marital.status`,sex,workclass,occupation,relationship,education
Resulted in 17127 records

Normalized Euclidean distance for (fnlwgt): 64.61
Hellinger Distance: 0.252

=====

Real			
	income	race	SUM(capital) COUNT(*)
1	<=50K	Asian-Pac-Islander	109724 1110
2	<=50K	Black	300347 4119
5	>50K	Amer-Indian-Eskimo	175710 55
0	<=50K	Amer-Indian-Eskimo	59843 415
3	<=50K	Other	8920 356

Synthetic			
	income	race	SUM(capital) COUNT(*)
1	<=50K	Asian-Pac-Islander	9441.715379 866.000000
2	<=50K	Black	17653.035310 3738.000000
5	>50K	Amer-Indian-Eskimo	-664.503375 14.000000
0	<=50K	Amer-Indian-Eskimo	-1788.391251 188.000000
3	<=50K	Other	778.285286 331.000000

SQL for Real:
SELECT income,race,SUM(capital), COUNT(*) FROM C1 GROUP BY income,race
Resulted in 10 records

SQL for Synthetic:
SELECT income,race,SUM(capital), COUNT(*) FROM C1_syn_06 GROUP BY income,race
Resulted in 10 records

Normalized Euclidean distance for (capital): 3.16
Hellinger Distance: 0.145

=====

Real											
	occupation	relationship	sex	marital.status	income	race	workclass	education	native.country	MIN(fnlwgt)	COUNT(*)
3755	Craft-repair	Not-in-family	Female	Married-spouse-absent	<=50K	White	Self-emp-inc	HS-grad	Puerto-Rico	nan	nan
6867	Farming-fishing	Husband	Male	Married-civ-spouse	<=50K	White	Self-emp-inc	Assoc-voc	Columbia	nan	nan
12129	Prof-specialty	Not-in-family	Female	Never-married	<=50K	White	Private	Prof-school	Puerto-Rico	nan	nan
3503	Craft-repair	Husband	Male	Married-civ-spouse	<=50K	White	Private	Some-college	Thailand	nan	nan
9099	Machine-op-inspct	Unmarried	Male	Never-married	<=50K	Black	Private	HS-grad	Peru	nan	nan

Synthetic											
	occupation	relationship	sex	marital.status	income	race	workclass	education	native.country	MIN(fnlwgt)	COUNT(*)
3755	Craft-repair	Not-in-family	Female	Married-spouse-absent	<=50K	White	Self-emp-inc	HS-grad	Puerto-Rico	175678.876887	1
6867	Farming-fishing	Husband	Male	Married-civ-spouse	<=50K	White	Self-emp-inc	Assoc-voc	Columbia	177651.128946	1
12129	Prof-specialty	Not-in-family	Female	Never-married	<=50K	White	Private	Prof-school	Puerto-Rico	165976.114730	3
3503	Craft-repair	Husband	Male	Married-civ-spouse	<=50K	White	Private	Some-college	Thailand	178421.128439	1
9099	Machine-op-inspct	Unmarried	Male	Never-married	<=50K	Black	Private	HS-grad	Peru	201688.943292	1

SQL for Real:
SELECT occupation,relationship,sex,`marital.status`,income,race,workclass,education,`native.country`,MIN(fnlwgt), COUNT(*) FROM C1 GROUP BY occupation,relationship,sex,`marital.status`,income,race,workclass,education,`native.country`
Resulted in 12664 records

SQL for Synthetic:
SELECT occupation,relationship,sex,`marital.status`,income,race,workclass,education,`native.country`,MIN(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY occupation,relationship,sex,`marital.status`,income,race,workclass,education,`native.country`
Resulted in 17127 records

Normalized Euclidean distance for (fnlwgt): 64.61

Hellinger Distance: 0.252

=====

Real											
	native.country	education	income	workclass	occupation	race	marital.status	relationship	AVG(capital)	COUNT(*)	
5104	Outlying-US(Guam-USVI-etc)	Some-college	<=50K	Private	Adm-clerical	White	Divorced	Other-relative	nan	nan	
13784	United-States	HS-grad	>50K	Self-emp-not-inc	Craft-repair	White	Never-married	Not-in-family	22868.600000	5.000000	
12368	United-States	Bachelors	<=50K	State-gov	Prof-specialty	Black	Never-married	Unmarried	0.000000	1.000000	
12625	United-States	Doctorate	>50K	Self-emp-inc	Sales	White	Married-civ-spouse	Husband	nan	nan	
8989	Puerto-Rico	Some-college	<=50K	Self-emp-not-inc	Prof-specialty	White	Divorced	Unmarried	nan	nan	

Synthetic											
	native.country	education	income	workclass	occupation	race	marital.status	relationship	AVG(capital)	COUNT(*)	
5104	Outlying-US(Guam-USVI-etc)	Some-college	<=50K	Private	Adm-clerical	White	Divorced	Other-relative	213.469716	1	
13784	United-States	HS-grad	>50K	Self-emp-not-inc	Craft-repair	White	Never-married	Not-in-family	-190.680263	1	
12368	United-States	Bachelors	<=50K	State-gov	Prof-specialty	Black	Never-married	Unmarried	-78.196953	1	
12625	United-States	Doctorate	>50K	Self-emp-inc	Sales	White	Married-civ-spouse	Husband	-180.439767	1	
8989	Puerto-Rico	Some-college	<=50K	Self-emp-not-inc	Prof-specialty	White	Divorced	Unmarried	23.109449	1	

SQL for Real:
SELECT `native.country`,education,income,workclass,occupation,race,`marital.status`,relationship,AVG(capital), COUNT(*) FROM C1 GROUP BY `native.country`,education,income,workclass,occupation,race,`marital.status`,relationship
Resulted in 11324 records

SQL for Synthetic:
SELECT `native.country`,education,income,workclass,occupation,race,`marital.status`,relationship,AVG(capital), COUNT(*) FROM C1_syn_06 GROUP BY `native.country`,education,income,workclass,occupation,race,`marital.status`,relationship
Resulted in 15410 records

Normalized Euclidean distance for (capital): 62.14

Hellinger Distance: 0.254

=====

Real									
	race	income	education	native.country	marital.status	occupation	workclass	MIN(age)	COUNT(*)
7603	White	<=50K	HS-grad	Laos	Married-civ-spouse	Sales	Self-emp-inc	nan	nan
9406	White	<=50K	Some-college	Haiti	Widowed	?	?	nan	nan
5301	White	<=50K	Assoc-acdm	United-States	Never-married	Sales	Private	20.000000	45.000000
1698	Black	<=50K	Bachelors	United-States	Married-civ-spouse	Sales	Private	24.000000	3.000000
2785	Black	>50K	HS-grad	Puerto-Rico	Married-civ-spouse	Machine-op-inspct	Private	nan	nan

Synthetic									
	race	income	education	native.country	marital.status	occupation	workclass	MIN(age)	COUNT(*)
7603	White	<=50K	HS-grad	Laos	Married-civ-spouse	Sales	Self-emp-inc	26	2
9406	White	<=50K	Some-college	Haiti	Widowed	?	?	59	1
5301	White	<=50K	Assoc-acdm	United-States	Never-married	Sales	Private	20	48
1698	Black	<=50K	Bachelors	United-States	Married-civ-spouse	Sales	Private	30	3
2785	Black	>50K	HS-grad	Puerto-Rico	Married-civ-spouse	Machine-op-inspct	Private	46	2

SQL for Real:
SELECT race,income,education,`native.country`,`marital.status`,occupation,workclass,MIN(age), COUNT(*) FROM C1 GROUP BY race,income,education,`native.country`,`marital.status`,occupation,workclass
Resulted in 8403 records

SQL for Synthetic:
SELECT race,income,education,`native.country`,`marital.status`,occupation,workclass,MIN(age), COUNT(*) FROM C1_syn_06 GROUP BY race,income,education,`native.country`,`marital.status`,occupation,workclass

Resulted in 11619 records

Normalized Euclidean distance for (age): 55.59

Hellinger Distance: 0.265

=====

Real					
	native.country	relationship	race	SUM(fnlwgt)	COUNT(*)
316	Laos	Own-child	Asian-Pac-Islander	875970.000000	3.000000
125	El-Salvador	Unmarried	White	5326863.000000	20.000000
484	Thailand	Other-relative	White	nan	nan
77	Columbia	Unmarried	Asian-Pac-Islander	nan	nan
86	Cuba	Own-child	White	4133709.000000	17.000000

Synthetic					
	native.country	relationship	race	SUM(fnlwgt)	COUNT(*)
316	Laos	Own-child	Asian-Pac-Islander	180075.427946	1
125	El-Salvador	Unmarried	White	936385.271671	5
484	Thailand	Other-relative	White	1244313.824766	7
77	Columbia	Unmarried	Asian-Pac-Islander	532638.467094	3
86	Cuba	Own-child	White	179577.687682	1

SQL for Real:
SELECT `native.country`,`relationship`,race,SUM(fnlwgt), COUNT(*) FROM C1 GROUP BY `native.country`,`relationship`,race

Resulted in 455 records

SQL for Synthetic:
SELECT `native.country`,`relationship`,race,SUM(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY `native.country`,`relationship`,race

Resulted in 549 records

Normalized Euclidean distance for (fnlwgt): 17.41

Hellinger Distance: 0.397

=====

Real							
	occupation	income	workclass	sex	education	SUM(capital)	COUNT(*)
1386	Prof-specialty	>50K	Self-emp-not-inc	Female	Assoc-acdm	0	3
1174	Other-service	>50K	State-gov	Female	Some-college	0	1
491	Exec-managerial	<=50K	Self-emp-inc	Female	Bachelors	0	7
620	Exec-managerial	>50K	Self-emp-inc	Male	HS-grad	211971	78
1476	Protective-serv	<=50K	State-gov	Male	Assoc-acdm	-1740	6

Synthetic							
	occupation	income	workclass	sex	education	SUM(capital)	COUNT(*)
1386	Prof-specialty	>50K	Self-emp-not-inc	Female	Assoc-acdm	-64.253621	1.000000
1174	Other-service	>50K	State-gov	Female	Some-college	nan	nan
491	Exec-managerial	<=50K	Self-emp-inc	Female	Bachelors	530.715776	12.000000
620	Exec-managerial	>50K	Self-emp-inc	Male	HS-grad	-738.609213	31.000000
1476	Protective-serv	<=50K	State-gov	Male	Assoc-acdm	227.498803	10.000000

SQL for Real:
SELECT occupation,income,workclass,sex,education,SUM(capital), COUNT(*) FROM C1 GROUP BY occupation,income,workclass,sex,education

Resulted in 1951 records

SQL for Synthetic:
SELECT occupation,income,workclass,sex,education,SUM(capital), COUNT(*) FROM C1_syn_06 GROUP BY occupation,income,workclass,sex,education

Resulted in 1896 records

Normalized Euclidean distance for (capital): 38.44

Hellinger Distance: 0.17

=====

Real				
	workclass	occupation	SUM(age)	COUNT(*)
44	Private	Farming-fishing	23582.000000	670.000000
45	Private	Handlers-cleaners	62151.000000	1923.000000
73	Self-emp-not-inc	Machine-op-inspct	2546.000000	59.000000
16	Federal-gov	Exec-managerial	11796.000000	268.000000
55	Self-emp-inc	Craft-repair	6964.000000	167.000000

Synthetic				
	workclass	occupation	SUM(age)	COUNT(*)
44	Private	Farming-fishing	19015	521
45	Private	Handlers-cleaners	61905	1926
73	Self-emp-not-inc	Machine-op-inspct	2879	64
16	Federal-gov	Exec-managerial	10320	225
55	Self-emp-inc	Craft-repair	6809	159

SQL for Real:
SELECT workclass,occupation,SUM(age), COUNT(*) FROM C1 GROUP BY workclass,occupation

Resulted in 84 records

SQL for Synthetic:
SELECT workclass,occupation,SUM(age), COUNT(*) FROM C1_syn_06 GROUP BY workclass,occupation
Resulted in 95 records

Normalized Euclidean distance for (age): 9.17

Hellinger Distance: 0.036

=====

Real							
	education	occupation	relationship	income	marital.status	SUM(capital)	COUNT(*)
950	7th-8th	Sales	Not-in-family	<=50K	Divorced	0	4
1955	Doctorate	Prof-specialty	Unmarried	<=50K	Married-spouse-absent	0	2
2145	HS-grad	Farming-fishing	Unmarried	<=50K	Divorced	0	7
977	9th	?	Not-in-family	<=50K	Never-married	0	3
2033	HS-grad	Adm-clerical	Own-child	<=50K	Married-civ-spouse	0	8

Synthetic							
	education	occupation	relationship	income	marital.status	SUM(capital)	COUNT(*)
950	7th-8th	Sales	Not-in-family	<=50K	Divorced	-171.448478	2.000000
1955	Doctorate	Prof-specialty	Unmarried	<=50K	Married-spouse-absent	25.313100	1.000000
2145	HS-grad	Farming-fishing	Unmarried	<=50K	Divorced	-150.244177	11.000000
977	9th	?	Not-in-family	<=50K	Never-married	-55.122475	3.000000
2033	HS-grad	Adm-clerical	Own-child	<=50K	Married-civ-spouse	152.155400	9.000000

SQL for Real:
SELECT education,occupation,relationship,income,`marital.status`,SUM(capital), COUNT(*) FROM C1 GROUP BY education,occupation,relationship,income,`marital.status`
Resulted in 3111 records

SQL for Synthetic:
SELECT education,occupation,relationship,income,`marital.status`,SUM(capital), COUNT(*) FROM C1_syn_06 GROUP BY education,occupation,relationship,income,`marital.status`
Resulted in 2930 records

Normalized Euclidean distance for (capital): 46.53

Hellinger Distance: 0.181

=====

Real							
	marital.status	relationship	native.country	occupation	education	MIN(fnlwgt)	COUNT(*)
5722	Never-married	Own-child	Canada	?	Assoc-voc	nan	nan
3508	Married-civ-spouse	Wife	Hungary	Prof-specialty	HS-grad	nan	nan
814	Divorced	Other-relative	United-States	Sales	Bachelors	198774.000000	2.000000
3432	Married-civ-spouse	Wife	Columbia	Other-service	Some-college	nan	nan
1763	Married-civ-spouse	Husband	Canada	Sales	Bachelors	204501.000000	3.000000

Synthetic							
	marital.status	relationship	native.country	occupation	education	MIN(fnlwgt)	COUNT(*)
5722	Never-married	Own-child	Canada	?	Assoc-voc	182697.599682	1
3508	Married-civ-spouse	Wife	Hungary	Prof-specialty	HS-grad	180141.940732	2
814	Divorced	Other-relative	United-States	Sales	Bachelors	181875.601920	1
3432	Married-civ-spouse	Wife	Columbia	Other-service	Some-college	177541.121928	1
1763	Married-civ-spouse	Husband	Canada	Sales	Bachelors	198795.272265	2

SQL for Real:
SELECT `marital.status`,relationship,`native.country`,occupation,education,MIN(fnlwgt), COUNT(*) FROM C1 GROUP BY `marital.status`,relationship,`native.country`,occupation,education
Resulted in 5498 records

SQL for Synthetic:
SELECT `marital.status`,relationship,`native.country`,occupation,education,MIN(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY `marital.status`,relationship,`native.country`,occupation,education
Resulted in 8263 records

Normalized Euclidean distance for (fnlwgt): 48.62

Hellinger Distance: 0.263

=====

Real				
	relationship	workclass	MIN(fnlwgt)	COUNT(*)
16	Other-relative	Local-gov	33386	64
13	Not-in-family	State-gov	19395	586
25	Own-child	Self-emp-inc	29582	93
26	Own-child	Self-emp-not-inc	25631	244
30	Unmarried	Local-gov	20101	436

Synthetic				
	relationship	workclass	MIN(fnlwgt)	COUNT(*)
16	Other-relative	Local-gov	148460.841202	87.000000
13	Not-in-family	State-gov	121336.094168	606.000000
25	Own-child	Self-emp-inc	147566.881431	118.000000
26	Own-child	Self-emp-not-inc	147338.179666	285.000000
30	Unmarried	Local-gov	141754.821183	407.000000

SQL for Real:
SELECT relationship,workclass,MIN(fnlwgt), COUNT(*) FROM C1 GROUP BY relationship,workclass
Resulted in 42 records

SQL for Synthetic:
SELECT relationship,workclass,MIN(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY relationship,workclass
Resulted in 42 records

Normalized Euclidean distance for (fnlwgt): 6.48
Hellinger Distance: 0.02

=====

Real				
	race	workclass	SUM(capital)	COUNT(*)
14	Black	?	66516	330
19	Black	Self-emp-not-inc	236814	136
12	Asian-Pac-Islander	Self-emp-not-inc	206528	100
18	Black	Self-emp-inc	153632	38
33	White	Self-emp-not-inc	5902959	3576

Synthetic				
	race	workclass	SUM(capital)	COUNT(*)
14	Black	?	-1239.763313	262.000000
19	Black	Self-emp-not-inc	1065.757357	93.000000
12	Asian-Pac-Islander	Self-emp-not-inc	1855.438249	82.000000
18	Black	Self-emp-inc	348.514647	36.000000
33	White	Self-emp-not-inc	-6677.688755	3618.000000

SQL for Real:
SELECT race,workclass,SUM(capital), COUNT(*) FROM C1 GROUP BY race,workclass
Resulted in 35 records

SQL for Synthetic:
SELECT race,workclass,SUM(capital), COUNT(*) FROM C1_syn_06 GROUP BY race,workclass
Resulted in 34 records

Normalized Euclidean distance for (capital): 5.83
Hellinger Distance: 0.044

=====

Real					
	workclass	relationship	native.country	MIN(capital)	COUNT(*)
364	Private	Other-relative	Holand-Netherlands	-2205	1
626	Self-emp-not-inc	Not-in-family	Mexico	0	7
100	?	Wife	Germany	0	1
557	Self-emp-inc	Unmarried	Canada	0	3
368	Private	Other-relative	India	0	8

Synthetic					
	workclass	relationship	native.country	MIN(capital)	COUNT(*)
364	Private	Other-relative	Holand-Netherlands	-192.221912	10.000000
626	Self-emp-not-inc	Not-in-family	Mexico	nan	nan
100	?	Wife	Germany	nan	nan
557	Self-emp-inc	Unmarried	Canada	nan	nan
368	Private	Other-relative	India	-174.164756	4.000000

SQL for Real:
SELECT workclass,relationship,`native.country`,MIN(capital), COUNT(*) FROM C1 GROUP BY workclass,relationship,`native.country`
Resulted in 751 records

SQL for Synthetic:
SELECT workclass,relationship,`native.country`,MIN(capital), COUNT(*) FROM C1_syn_06 GROUP BY workclass,relationship,`native.country`
Resulted in 732 records

Normalized Euclidean distance for (capital): 20.37
Hellinger Distance: 0.387

=====

Real											
	education	relationship	income	native.country	marital.status	race	occupation	workclass	sex	MAX(capital)	COUNT(*)
15655	Some-college	Other-relative	<=50K	United-States	Never-married	White	Prof-specialty	Local-gov	Male	0.000000	2.000000
7660	Bachelors	Wife	>50K	?	Married-civ-spouse	Asian-Pac-Islander	Adm-clerical	Private	Female	nan	nan
9826	HS-grad	Not-in-family	<=50K	United-States	Married-spouse-absent	White	Tech-support	Private	Female	0.000000	1.000000
3746	Assoc-acdm	Not-in-family	<=50K	United-States	Never-married	White	Tech-support	Private	Male	0.000000	10.000000
2892	9th	Not-in-family	<=50K	?	Divorced	White	Other-service	Private	Female	nan	nan

Synthetic											
	education	relationship	income	native.country	marital.status	race	occupation	workclass	sex	MAX(capital)	COUNT(*)
15655	Some-college	Other-relative	<=50K	United-States	Never-married	White	Prof-specialty	Local-gov	Male	54.360139	1
7660	Bachelors	Wife	>50K	?	Married-civ-spouse	Asian-Pac-Islander	Adm-clerical	Private	Female	-225.759230	1
9826	HS-grad	Not-in-family	<=50K	United-States	Married-spouse-absent	White	Tech-support	Private	Female	63.025435	1
3746	Assoc-acdm	Not-in-family	<=50K	United-States	Never-married	White	Tech-support	Private	Male	9.491772	2
2892	9th	Not-in-family	<=50K	?	Divorced	White	Other-service	Private	Female	-159.955537	1

SQL for Real:

SELECT education,relationship,income,`native.country`,`marital.status`,race,occupation,workclass,sex,MAX(capital), COUNT(*) FROM C1 GROUP BY education,relationship,income,`native.country`,`marital.status`,race,occupation,workclass,sex

Resulted in 12664 records

SQL for Synthetic:

SELECT education,relationship,income,`native.country`,`marital.status`,race,occupation,workclass,sex,MAX(capital), COUNT(*) FROM C1_syn_06 GROUP BY education,relationship,income,`native.country`,`marital.status`,race,occupation,workclass,sex

Resulted in 17127 records

Normalized Euclidean distance for (capital): 64.61

Hellinger Distance: 0.252

=====

Real											
	sex	native.country	income	education	occupation	marital.status	race	workclass	relationship	AVG(fnlwgt)	COUNT(*)
16636	Male	United-States	>50K	Assoc-acdm	Armed-Forces	Never-married	White	Federal-gov	Own-child	nan	nan
4382	Female	Scotland	<=50K	Assoc-acdm	Other-service	Married-civ-spouse	White	Local-gov	Wife	nan	nan
7903	Male	?	<=50K	Masters	Exec-managerial	Married-civ-spouse	Black	Private	Husband	103408.000000	1.000000
9988	Male	Japan	<=50K	HS-grad	Sales	Divorced	White	Self-emp-inc	Unmarried	nan	nan
6882	Female	United-States	<=50K	Some-college	Adm-clerical	Separated	White	Private	Other-relative	32958.000000	1.000000

Synthetic											
	sex	native.country	income	education	occupation	marital.status	race	workclass	relationship	AVG(fnlwgt)	COUNT(*)
16636	Male	United-States	>50K	Assoc-acdm	Armed-Forces	Never-married	White	Federal-gov	Own-child	193422.466221	1
4382	Female	Scotland	<=50K	Assoc-acdm	Other-service	Married-civ-spouse	White	Local-gov	Wife	180030.760019	1
7903	Male	?	<=50K	Masters	Exec-managerial	Married-civ-spouse	Black	Private	Husband	184700.865109	1
9988	Male	Japan	<=50K	HS-grad	Sales	Divorced	White	Self-emp-inc	Unmarried	178176.562640	1
6882	Female	United-States	<=50K	Some-college	Adm-clerical	Separated	White	Private	Other-relative	181712.006559	1

SQL for Real:

SELECT sex,`native.country`,income,education,occupation,`marital.status`,race,workclass,relationship,AVG(fnlwgt), COUNT(*) FROM C1 GROUP BY sex,`native.country`,income,education,occupation,`marital.status`,race,workclass,relationship

Resulted in 12664 records

SQL for Synthetic:

SELECT sex,`native.country`,income,education,occupation,`marital.status`,race,workclass,relationship,AVG(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY sex,`native.country`,income,education,occupation,`marital.status`,race,workclass,relationship

Resulted in 17127 records

Normalized Euclidean distance for (fnlwgt): 64.61

Hellinger Distance: 0.252

=====

Real											
	marital.status	sex	race	relationship	workclass	native.country	education	income	SUM(fnlwgt)	COUNT(*)	
2219	Married-civ-spouse	Female	White	Wife	?	Taiwan	Bachelors	<=50K	nan	nan	
8657	Widowed	Female	White	Not-in-family	?	United-States	Bachelors	<=50K	1217251.000000	7.000000	
6837	Never-married	Male	White	Not-in-family	?	United-States	Some-college	<=50K	6981532.000000	33.000000	
5976	Never-married	Female	White	Own-child	Private	Iran	Bachelors	<=50K	nan	nan	
5864	Never-married	Female	White	Own-child	Federal-gov	Peru	Some-college	<=50K	nan	nan	

Synthetic											
	marital.status	sex	race	relationship	workclass	native.country	education	income	SUM(fnlwgt)	COUNT(*)	
2219	Married-civ-spouse	Female	White	Wife	?	Taiwan	Bachelors	<=50K	179913.334884	1	
8657	Widowed	Female	White	Not-in-family	?	United-States	Bachelors	<=50K	567783.717992	4	
6837	Never-married	Male	White	Not-in-family	?	United-States	Some-college	<=50K	4308684.928030	23	
5976	Never-married	Female	White	Own-child	Private	Iran	Bachelors	<=50K	180655.073062	1	
5864	Never-married	Female	White	Own-child	Federal-gov	Peru	Some-college	<=50K	204050.999748	1	

SQL for Real:

SELECT `marital.status`,sex,race,relationship,workclass,`native.country`,education,income,SUM(fnlwgt), COUNT(*) FROM C1 GROUP BY `marital.status`,sex,race,relationship,workclass,`native.country`,education,income

Resulted in 6863 records

SQL for Synthetic:

SELECT `marital.status`,sex,race,relationship,workclass,`native.country`,education,income,SUM(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY `marital.status`,sex,race,relationship,workclass,`native.country`,education,income

Resulted in 9165 records

Normalized Euclidean distance for (fnlwgt): 50.11

Hellinger Distance: 0.311

=====

Real					
	race	sex	income	SUM(` hours.per.week`)	COUNT(*)
7	Asian-Pac-Islander	Male	>50K	15507	340
4	Asian-Pac-Islander	Female	<=50K	16370	448
12	Other	Female	<=50K	5108	144
0	Amer-Indian-Eskimo	Female	<=50K	6236	170
6	Asian-Pac-Islander	Male	<=50K	25824	662

Synthetic					
	race	sex	income	SUM(` hours.per.week`)	COUNT(*)
7	Asian-Pac-Islander	Male	>50K	4451.509586	111.000000
4	Asian-Pac-Islander	Female	<=50K	12263.689885	307.000000
12	Other	Female	<=50K	6092.589092	153.000000
0	Amer-Indian-Eskimo	Female	<=50K	2553.226256	64.000000
6	Asian-Pac-Islander	Male	<=50K	22375.374734	559.000000

SQL for Real:
SELECT race,sex,income,SUM(` hours.per.week`), COUNT(*) FROM C1 GROUP BY race,sex,income

Resulted in 20 records

SQL for Synthetic:
SELECT race,sex,income,SUM(` hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY race,sex,income

Resulted in 20 records

Normalized Euclidean distance for (` hours.per.week`): 4.47

Hellinger Distance: 0.145

=====

Real							
	income	sex	native.country	education	marital.status	MAX(age)	COUNT(*)
2201	>50K	Male	Puerto-Rico	Bachelors	Never-married	nan	nan
2128	>50K	Male	Iran	Some-college	Married-civ-spouse	47.000000	2.000000
1765	<=50K	Male	Thailand	Some-college	Divorced	nan	nan
987	<=50K	Male	?	Assoc-acdm	Married-civ-spouse	42.000000	2.000000
466	<=50K	Female	Japan	Assoc-voc	Never-married	nan	nan

Synthetic							
	income	sex	native.country	education	marital.status	MAX(age)	COUNT(*)
2201	>50K	Male	Puerto-Rico	Bachelors	Never-married	47	5
2128	>50K	Male	Iran	Some-college	Married-civ-spouse	30	1
1765	<=50K	Male	Thailand	Some-college	Divorced	24	1
987	<=50K	Male	?	Assoc-acdm	Married-civ-spouse	64	14
466	<=50K	Female	Japan	Assoc-voc	Never-married	37	3

SQL for Real:
SELECT income,sex,` native.country`,education,` marital.status`,MAX(age), COUNT(*) FROM C1 GROUP BY income,sex,` native.country`,education,` marital.status`

Resulted in 2126 records

SQL for Synthetic:
SELECT income,sex,` native.country`,education,` marital.status`,MAX(age), COUNT(*) FROM C1_syn_06 GROUP BY income,sex,` native.country`,education,` marital.status`

Resulted in 2296 records

Normalized Euclidean distance for (age): 31.22

Hellinger Distance: 0.38

=====

Real				
	occupation	sex	SUM(` hours.per.week`)	COUNT(*)
0	?	Female	38062.000000	1273.000000
1	?	Male	51270.000000	1536.000000
23	Protective-serv	Male	37396.000000	861.000000
25	Sales	Male	157836.000000	3557.000000
27	Tech-support	Male	36301.000000	884.000000

Synthetic				
	occupation	sex	SUM(` hours.per.week`)	COUNT(*)
0	?	Female	54805.924489	1381
1	?	Male	57906.940697	1459
23	Protective-serv	Male	34958.469808	874
25	Sales	Male	142129.655037	3537
27	Tech-support	Male	35272.407887	882

SQL for Real:
SELECT occupation,sex,SUM(` hours.per.week`), COUNT(*) FROM C1 GROUP BY occupation,sex

Resulted in 29 records

SQL for Synthetic:
SELECT occupation,sex,SUM(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY occupation,sex
Resulted in 30 records

Normalized Euclidean distance for (`hours.per.week`): 5.39

Hellinger Distance: 0.028

=====

Real			
	relationship	MAX(capital)	COUNT(*)
5	Wife	99999	2331
3	Own-child	99999	7581
0	Husband	99999	19716
1	Not-in-family	99999	12583
4	Unmarried	99999	5125

Synthetic			
	relationship	MAX(capital)	COUNT(*)
5	Wife	659.125210	2479.000000
3	Own-child	439.857261	8227.000000
0	Husband	872.469211	19353.000000
1	Not-in-family	902.309265	12433.000000
4	Unmarried	464.951776	4999.000000

SQL for Real:
SELECT relationship,MAX(capital), COUNT(*) FROM C1 GROUP BY relationship
Resulted in 6 records

SQL for Synthetic:
SELECT relationship,MAX(capital), COUNT(*) FROM C1_syn_06 GROUP BY relationship
Resulted in 6 records

Normalized Euclidean distance for (capital): 2.45

Hellinger Distance: 0.015

=====

Real			
	income	MAX(capital)	COUNT(*)
0	<=50K	41310	37155
1	>50K	99999	11687

Synthetic			
	income	MAX(capital)	COUNT(*)
0	<=50K	902.309265	44080.000000
1	>50K	574.776427	4762.000000

SQL for Real:
SELECT income,MAX(capital), COUNT(*) FROM C1 GROUP BY income
Resulted in 2 records

SQL for Synthetic:
SELECT income,MAX(capital), COUNT(*) FROM C1_syn_06 GROUP BY income
Resulted in 2 records

Normalized Euclidean distance for (capital): 1.41

Hellinger Distance: 0.137

=====

Real											
	relationship	sex	income	marital.status	race	occupation	education	native.country	workclass	MIN(`hours.per.week`)	COUNT(*)
5280	Not-in-family	Female	<=50K	Never-married	Other	Adm-clerical	Assoc-voc	United-States	Private	35.000000	1.000000
12439	Own-child	Male	<=50K	Never-married	White	Exec-managerial	Some-college	Laos	Private	nan	nan
3186	Husband	Male	>50K	Married-civ-spouse	Asian-Pac-Islander	Exec-managerial	HS-grad	United-States	Self-emp-inc	40.000000	1.000000
10922	Own-child	Female	<=50K	Never-married	Other	Sales	Some-college	United-States	Private	20.000000	1.000000
6865	Not-in-family	Male	<=50K	Divorced	Black	Sales	HS-grad	United-States	Private	nan	nan
Synthetic											
	relationship	sex	income	marital.status	race	occupation	education	native.country	workclass	MIN(`hours.per.week`)	COUNT(*)
5280	Not-in-family	Female	<=50K	Never-married	Other	Adm-clerical	Assoc-voc	United-States	Private	40.020390	1
12439	Own-child	Male	<=50K	Never-married	White	Exec-managerial	Some-college	Laos	Private	39.981428	1
3186	Husband	Male	>50K	Married-civ-spouse	Asian-Pac-Islander	Exec-managerial	HS-grad	United-States	Self-emp-inc	40.355496	1
10922	Own-child	Female	<=50K	Never-married	Other	Sales	Some-college	United-States	Private	39.436983	2
6865	Not-in-family	Male	<=50K	Divorced	Black	Sales	HS-grad	United-States	Private	40.410807	1

SQL for Real:
SELECT relationship,sex,income,`marital.status`,race,occupation,education,`native.country`,workclass,MIN(`hours.per.week`), COUNT(*) FROM C1 GROUP BY relationship,sex,income,`marital.status`,race,occupation,education,`native.country`,workclass

Resulted in 12664 records

```
SQL for Synthetic:
SELECT relationship,sex,income,`marital.status`,race,occupation,education,`native.country`,workclass,MIN(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY
relationship,sex,income,`marital.status`,race,occupation,education,`native.country`,workclass
```

Resulted in 17127 records

Normalized Euclidean distance for (`hours.per.week`): 64.61

Hellinger Distance: 0.252

=====

Real					
	income	sex	occupation	SUM(fnlwgt)	COUNT(*)
13	<=50K	Female	Transport-moving	22414590	114
1	<=50K	Female	Adm-clerical	653797659	3460
3	<=50K	Female	Exec-managerial	244876970	1327
26	<=50K	Male	Sales	422850837	2216
15	<=50K	Male	Adm-clerical	278084013	1383

Synthetic					
	income	sex	occupation	SUM(fnlwgt)	COUNT(*)
13	<=50K	Female	Transport-moving	20329849.218387	113.000000
1	<=50K	Female	Adm-clerical	676108083.870327	3717.000000
3	<=50K	Female	Exec-managerial	309798992.128707	1714.000000
26	<=50K	Male	Sales	555375990.992196	3044.000000
15	<=50K	Male	Adm-clerical	304251917.873978	1679.000000

```
SQL for Real:
SELECT income,sex,occupation,SUM(fnlwgt), COUNT(*) FROM C1 GROUP BY income,sex,occupation
```

Resulted in 57 records

```
SQL for Synthetic:
SELECT income,sex,occupation,SUM(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY income,sex,occupation
```

Resulted in 57 records

Normalized Euclidean distance for (fnlwgt): 7.48

Hellinger Distance: 0.148

=====

Real											
	sex	native.country	income	marital.status	occupation	relationship	race	education	workclass	MAX(`hours.per.week`)	COUNT(*)
10278	Male	Laos	<=50K	Married-civ-spouse	Other-service	Husband	White	HS-grad	Self-emp-not-inc	nan	nan
585	Female	Columbia	<=50K	Married-civ-spouse	Adm-clerical	Wife	White	HS-grad	Private	nan	nan
3687	Female	Puerto-Rico	<=50K	Never-married	Other-service	Other-relative	Black	11th	Private	nan	nan
10293	Male	Laos	<=50K	Married-civ-spouse	Protective-serv	Husband	White	Bachelors	Private	nan	nan
136	Female	?	<=50K	Married-civ-spouse	Prof-specialty	Other-relative	White	Masters	Local-gov	nan	nan

Synthetic											
	sex	native.country	income	marital.status	occupation	relationship	race	education	workclass	MAX(`hours.per.week`)	COUNT(*)
10278	Male	Laos	<=50K	Married-civ-spouse	Other-service	Husband	White	HS-grad	Self-emp-not-inc	40.251916	1
585	Female	Columbia	<=50K	Married-civ-spouse	Adm-clerical	Wife	White	HS-grad	Private	40.092411	4
3687	Female	Puerto-Rico	<=50K	Never-married	Other-service	Other-relative	Black	11th	Private	38.985975	1
10293	Male	Laos	<=50K	Married-civ-spouse	Protective-serv	Husband	White	Bachelors	Private	39.990177	1
136	Female	?	<=50K	Married-civ-spouse	Prof-specialty	Other-relative	White	Masters	Local-gov	39.998828	1

```
SQL for Real:
SELECT sex,`native.country`,income,`marital.status`,occupation,relationship,race,education,workclass,MAX(`hours.per.week`), COUNT(*) FROM C1 GROUP BY
sex,`native.country`,income,`marital.status`,occupation,relationship,race,education,workclass
```

Resulted in 12664 records

```
SQL for Synthetic:
SELECT sex,`native.country`,income,`marital.status`,occupation,relationship,race,education,workclass,MAX(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY
sex,`native.country`,income,`marital.status`,occupation,relationship,race,education,workclass
```

Resulted in 17127 records

Normalized Euclidean distance for (`hours.per.week`): 64.61

Hellinger Distance: 0.252

=====

Real								
	race	native.country	income	marital.status	relationship	workclass	MIN(capital)	COUNT(*)
2059	White	Nicaragua	<=50K	Never-married	Own-child	Private	0.000000	8.000000
346	Asian-Pac-Islander	United-States	<=50K	Divorced	Unmarried	Federal-gov	0.000000	2.000000
1403	White	Columbia	<=50K	Married-civ-spouse	Other-relative	Private	0.000000	1.000000
1223	White	?	<=50K	Divorced	Unmarried	Local-gov	0.000000	1.000000
1859	White	Jamaica	<=50K	Widowed	Own-child	Private	nan	nan

Synthetic								
	race	native.country	income	marital.status	relationship	workclass	MIN(capital)	COUNT(*)
2059	White	Nicaragua	<=50K	Never-married	Own-child	Private	-170.650011	68
346	Asian-Pac-Islander	United-States	<=50K	Divorced	Unmarried	Federal-gov	103.320185	1
1403	White	Columbia	<=50K	Married-civ-spouse	Other-relative	Private	-268.506274	2
1223	White	?	<=50K	Divorced	Unmarried	Local-gov	-117.326702	3
1859	White	Jamaica	<=50K	Widowed	Own-child	Private	-43.644932	1

SQL for Real:
SELECT race,`native.country`,income,`marital.status`,relationship,workclass,MIN(capital), COUNT(*) FROM C1 GROUP BY race,`native.country`,income,`marital.status`,relationship,workclass

Resulted in 2280 records

SQL for Synthetic:
SELECT race,`native.country`,income,`marital.status`,relationship,workclass,MIN(capital), COUNT(*) FROM C1_syn_06 GROUP BY race,`native.country`,income,`marital.status`,relationship,workclass

Resulted in 2837 records

Normalized Euclidean distance for (capital): 31.26

Hellinger Distance: 0.384

=====

Real					
	race	relationship	marital.status	AVG(`hours.per.week`)	COUNT(*)
34	Asian-Pac-Islander	Own-child	Married-civ-spouse	39.000000	12.000000
85	Other	Not-in-family	Widowed	38.000000	4.000000
29	Asian-Pac-Islander	Other-relative	Married-civ-spouse	41.600000	30.000000
99	Other	Wife	Married-civ-spouse	35.869565	23.000000
98	Other	Unmarried	Widowed	37.000000	5.000000

Synthetic					
	race	relationship	marital.status	AVG(`hours.per.week`)	COUNT(*)
34	Asian-Pac-Islander	Own-child	Married-civ-spouse	40.013544	5
85	Other	Not-in-family	Widowed	39.719706	4
29	Asian-Pac-Islander	Other-relative	Married-civ-spouse	39.918964	11
99	Other	Wife	Married-civ-spouse	39.954263	20
98	Other	Unmarried	Widowed	40.009190	3

SQL for Real:
SELECT race,relationship,`marital.status`,AVG(`hours.per.week`), COUNT(*) FROM C1 GROUP BY race,relationship,`marital.status`

Resulted in 125 records

SQL for Synthetic:
SELECT race,relationship,`marital.status`,AVG(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY race,relationship,`marital.status`

Resulted in 140 records

Normalized Euclidean distance for (`hours.per.week`): 10.63

Hellinger Distance: 0.052

=====

Real			
	workclass	MAX(age)	COUNT(*)
4	Self-emp-inc	85	1695
3	Private	90	33906
6	State-gov	81	1981
1	Federal-gov	90	1432
0	?	90	2830

Synthetic			
	workclass	MAX(age)	COUNT(*)
4	Self-emp-inc	84.000000	1645.000000
3	Private	90.000000	34022.000000
6	State-gov	81.000000	1982.000000
1	Federal-gov	79.000000	1426.000000
0	?	90.000000	2844.000000

SQL for Real:
SELECT workclass,MAX(age), COUNT(*) FROM C1 GROUP BY workclass

Resulted in 7 records

SQL for Synthetic:
SELECT workclass,MAX(age), COUNT(*) FROM C1_syn_06 GROUP BY workclass

Resulted in 7 records

Normalized Euclidean distance for (age): 2.65

Hellinger Distance: 0.003

=====

Real									
	relationship	race	education	sex	occupation	workclass	income	MAX(capital)	COUNT(*)
1379	Husband	White	Doctorate	Male	Protective-serv	Local-gov	>50K	0	1
1946	Not-in-family	Asian-Pac-Islander	Bachelors	Male	Prof-specialty	Federal-gov	<=50K	0	1
2739	Not-in-family	White	7th-8th	Male	Machine-op-inspct	Private	<=50K	0	9
504	Husband	Black	HS-grad	Male	Adm-clerical	Self-emp-not-inc	<=50K	0	1
3544	Not-in-family	White	Some-college	Male	Farming-fishing	Local-gov	<=50K	0	2

Synthetic									
	relationship	race	education	sex	occupation	workclass	income	MAX(capital)	COUNT(*)
1379	Husband	White	Doctorate	Male	Protective-serv	Local-gov	>50K	nan	nan
1946	Not-in-family	Asian-Pac-Islander	Bachelors	Male	Prof-specialty	Federal-gov	<=50K	nan	nan
2739	Not-in-family	White	7th-8th	Male	Machine-op-inspct	Private	<=50K	50.279206	4.000000
504	Husband	Black	HS-grad	Male	Adm-clerical	Self-emp-not-inc	<=50K	nan	nan
3544	Not-in-family	White	Some-college	Male	Farming-fishing	Local-gov	<=50K	183.240194	2.000000

SQL for Real:
SELECT relationship,race,education,sex,occupation,workclass,income,MAX(capital), COUNT(*) FROM C1 GROUP BY relationship,race,education,sex,occupation,workclass,income

Resulted in 7027 records

SQL for Synthetic:
SELECT relationship,race,education,sex,occupation,workclass,income,MAX(capital), COUNT(*) FROM C1_syn_06 GROUP BY relationship,race,education,sex,occupation,workclass,income

Resulted in 6287 records

Normalized Euclidean distance for (capital): 62.94

Hellinger Distance: 0.201

=====

Real											
	income	marital.status	occupation	sex	native.country	race	workclass	education	relationship	MAX(`hours.per.week`)	COUNT(*)
6396	<=50K	Married-civ-spouse	Protective-serv	Female	Puerto-Rico	White	Private	HS-grad	Wife	nan	nan
8476	<=50K	Never-married	Adm-clerical	Female	Puerto-Rico	White	Private	Bachelors	Own-child	nan	nan
2846	<=50K	Divorced	Tech-support	Female	Peru	White	Local-gov	Assoc-acdm	Not-in-family	nan	nan
7803	<=50K	Married-spouse-absent	Sales	Male	United-States	Asian-Pac-Islander	Self-emp-not-inc	Masters	Own-child	nan	nan
5594	<=50K	Married-civ-spouse	Other-service	Male	?	White	Private	HS-grad	Husband	45.000000	8.000000

Synthetic											
	income	marital.status	occupation	sex	native.country	race	workclass	education	relationship	MAX(`hours.per.week`)	COUNT(*)
6396	<=50K	Married-civ-spouse	Protective-serv	Female	Puerto-Rico	White	Private	HS-grad	Wife	39.944647	1
8476	<=50K	Never-married	Adm-clerical	Female	Puerto-Rico	White	Private	Bachelors	Own-child	40.009431	5
2846	<=50K	Divorced	Tech-support	Female	Peru	White	Local-gov	Assoc-acdm	Not-in-family	39.963949	1
7803	<=50K	Married-spouse-absent	Sales	Male	United-States	Asian-Pac-Islander	Self-emp-not-inc	Masters	Own-child	40.006911	1
5594	<=50K	Married-civ-spouse	Other-service	Male	?	White	Private	HS-grad	Husband	40.048950	10

SQL for Real:
SELECT income`,`marital.status`,`occupation,sex`,`native.country`,`race,workclass,education,relationship,MAX(`hours.per.week`), COUNT(*) FROM C1 GROUP BY income`,`marital.status`,`occupation,sex`,`native.country`,`race,workclass,education,relationship

Resulted in 12664 records

SQL for Synthetic:
SELECT income`,`marital.status`,`occupation,sex`,`native.country`,`race,workclass,education,relationship,MAX(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY income`,`marital.status`,`occupation,sex`,`native.country`,`race,workclass,education,relationship

Resulted in 17127 records

Normalized Euclidean distance for (`hours.per.week`): 64.61

Hellinger Distance: 0.252

=====

Real										
	workclass	sex	occupation	race	relationship	education	income	MIN(fnlwgt)	COUNT(*)	
194	?	Male	?	Amer-Indian-Eskimo	Unmarried	Some-college	<=50K	98145	2	
1295	Local-gov	Male	Adm-clerical	Other	Not-in-family	Some-college	<=50K	205262	1	
5759	Self-emp-not-inc	Female	Sales	White	Unmarried	Assoc-acdm	>50K	128516	1	
6247	Self-emp-not-inc	Male	Sales	White	Husband	Assoc-voc	>50K	26145	6	
649	Federal-gov	Male	Armed-Forces	White	Husband	Assoc-voc	>50K	190653	1	

Synthetic										
	workclass	sex	occupation	race	relationship	education	income	MIN(fnlwgt)	COUNT(*)	
194	?	Male	?	Amer-Indian-Eskimo	Unmarried	Some-college	<=50K	nan	nan	
1295	Local-gov	Male	Adm-clerical	Other	Not-in-family	Some-college	<=50K	197979.302587	1.000000	
5759	Self-emp-not-inc	Female	Sales	White	Unmarried	Assoc-acdm	>50K	nan	nan	
6247	Self-emp-not-inc	Male	Sales	White	Husband	Assoc-voc	>50K	179600.881859	1.000000	
649	Federal-gov	Male	Armed-Forces	White	Husband	Assoc-voc	>50K	nan	nan	

SQL for Real:
SELECT workclass,sex,occupation,race,relationship,education,income,MIN(fnlwgt), COUNT(*) FROM C1 GROUP BY

workclass,sex,occupation,race,relationship,education,income

Resulted in 7027 records

SQL for Synthetic:

SELECT workclass,sex,occupation,race,relationship,education,income,MIN(fnlwgt), COUNT(*) FROM C1_syn_06 GROUP BY workclass,sex,occupation,race,relationship,education,income

Resulted in 6287 records

Normalized Euclidean distance for (fnlwgt): 62.94

Hellinger Distance: 0.201

=====

Real										
	education	race	occupation	relationship	workclass	sex	income	native.country	SUM(capital)	COUNT(*)
1173	11th	White	Sales	Other-relative	Private	Female	<=50K	Nicaragua	nan	nan
5905	Bachelors	White	Prof-specialty	Husband	Private	Male	>50K	Peru	-1848.000000	1.000000
1151	11th	White	Protective-serv	Own-child	Private	Female	<=50K	United-States	0.000000	2.000000
9307	HS-grad	White	Other-service	Unmarried	Private	Male	<=50K	Thailand	nan	nan
6039	Bachelors	White	Prof-specialty	Not-in-family	State-gov	Female	<=50K	Puerto-Rico	nan	nan

Synthetic										
	education	race	occupation	relationship	workclass	sex	income	native.country	SUM(capital)	COUNT(*)
1173	11th	White	Sales	Other-relative	Private	Female	<=50K	Nicaragua	-11.666286	1
5905	Bachelors	White	Prof-specialty	Husband	Private	Male	>50K	Peru	-154.563193	4
1151	11th	White	Protective-serv	Own-child	Private	Female	<=50K	United-States	114.923499	1
9307	HS-grad	White	Other-service	Unmarried	Private	Male	<=50K	Thailand	-122.335465	1
6039	Bachelors	White	Prof-specialty	Not-in-family	State-gov	Female	<=50K	Puerto-Rico	-83.164883	9

SQL for Real:

SELECT education,race,occupation,relationship,workclass,sex,income,`native.country`,SUM(capital), COUNT(*) FROM C1 GROUP BY education,race,occupation,relationship,workclass,sex,income,`native.country`

Resulted in 9992 records

SQL for Synthetic:

SELECT education,race,occupation,relationship,workclass,sex,income,`native.country`,SUM(capital), COUNT(*) FROM C1_syn_06 GROUP BY education,race,occupation,relationship,workclass,sex,income,`native.country`

Resulted in 13906 records

Normalized Euclidean distance for (capital): 60.46

Hellinger Distance: 0.257

=====

Real			
	education	SUM(age)	COUNT(*)
0	10th	52647	1389
1	11th	58241	1812
12	Masters	116926	2657
7	Assoc-acdm	60481	1601
8	Assoc-voc	79660	2061

Synthetic			
	education	SUM(age)	COUNT(*)
0	10th	54965.000000	1430.000000
1	11th	55417.000000	1694.000000
12	Masters	121790.000000	2810.000000
7	Assoc-acdm	65407.000000	1755.000000
8	Assoc-voc	73863.000000	1935.000000

SQL for Real:

SELECT education,SUM(age), COUNT(*) FROM C1 GROUP BY education

Resulted in 16 records

SQL for Synthetic:

SELECT education,SUM(age), COUNT(*) FROM C1_syn_06 GROUP BY education

Resulted in 16 records

Normalized Euclidean distance for (age): 4.0

Hellinger Distance: 0.026

=====

Real		
	sex	AVG(`hours.per.week`) COUNT(*)
0	Female	36.400692 16192
1	Male	42.416845 32650

Synthetic		
	sex	AVG(`hours.per.week`) COUNT(*)
0	Female	39.881125 17114.000000
1	Male	40.040656 31728.000000

SQL for Real:

SELECT sex,AVG(`hours.per.week`), COUNT(*) FROM C1 GROUP BY sex

Resulted in 2 records

SQL for Synthetic:
SELECT sex,AVG(`hours.per.week`), COUNT(*) FROM C1_syn_06 GROUP BY sex
Resulted in 2 records

Normalized Euclidean distance for (`hours.per.week`): 1.41

Hellinger Distance: 0.014

=====

Real										
	income	workclass	native.country	education	race	occupation	sex	relationship	MAX(age)	COUNT(*)
2079	<=50K	Local-gov	Scotland	Masters	White	Prof-specialty	Female	Wife	nan	nan
13434	>50K	Private	United-States	Some-college	White	Handlers-cleaners	Male	Husband	58.000000	14.000000
4806	<=50K	Private	Italy	Bachelors	White	Other-service	Male	Husband	nan	nan
10874	<=50K	Self-emp-not-inc	Thailand	Some-college	White	Other-service	Male	Not-in-family	nan	nan
13533	>50K	Self-emp-inc	Puerto-Rico	Some-college	White	Sales	Male	Husband	nan	nan

Synthetic										
	income	workclass	native.country	education	race	occupation	sex	relationship	MAX(age)	COUNT(*)
2079	<=50K	Local-gov	Scotland	Masters	White	Prof-specialty	Female	Wife	52	2
13434	>50K	Private	United-States	Some-college	White	Handlers-cleaners	Male	Husband	50	3
4806	<=50K	Private	Italy	Bachelors	White	Other-service	Male	Husband	37	1
10874	<=50K	Self-emp-not-inc	Thailand	Some-college	White	Other-service	Male	Not-in-family	24	1
13533	>50K	Self-emp-inc	Puerto-Rico	Some-college	White	Sales	Male	Husband	68	8

SQL for Real:
SELECT income,workclass,`native.country`,education,race,occupation,sex,relationship,MAX(age), COUNT(*) FROM C1 GROUP BY income,workclass,`native.country`,education,race,occupation,sex,relationship
Resulted in 9992 records

SQL for Synthetic:
SELECT income,workclass,`native.country`,education,race,occupation,sex,relationship,MAX(age), COUNT(*) FROM C1_syn_06 GROUP BY income,workclass,`native.country`,education,race,occupation,sex,relationship
Resulted in 13906 records

Normalized Euclidean distance for (age): 60.46

Hellinger Distance: 0.257

=====

Hellinger Distance Summary: {'mean': 0.2078041925712827, 'median': 0.2523067213216237, 'stddev': 0.12632558408604722}

Euclidean distance Summary : {'mean': 32.22667937857911, 'median': 29.2506405686264, 'stddev': 23.717106405087968}

=====