

adult_1700864246646

Real data sample:

	age	workclass	fnlwgt	education	education-num	marital-status	occupation	relationship	race	sex	capital-gain	capital-loss	hours-per-week	native-country	income
0	39	State-gov	77516	Bachelors	13	Never-married	Adm-clerical	Not-in-family	White	Male	2174	0	40	United-States	>50K
1	50	Self-emp-not-inc	83314	Bachelors	13	Married-civ-spouse	Exec-managerial	Husband	White	Male	0	0	13	United-States	>50K
2	38	Private	215646	HS-grad	9	Divorced	Handlers-cleaners	Not-in-family	White	Male	0	0	40	United-States	>50K
3	53	Private	234721	11th	7	Married-civ-spouse	Transportation-machine-op	Husband	Black	Male	0	0	40	United-States	>50K
4	28	Private	338409	Bachelors	13	Married-civ-spouse	Spec-proc	Wife	Black	Female	0	0	40	Cuba	<=50K

Real data stats:

	age	fnlwgt	education-num	capital-gain	capital-loss	hours-per-week
count	32561.000000	3.256100e+04	32561.000000	32561.000000	32561.000000	32561.000000
mean	38.581647	1.897784e+05	10.080679	1077.648844	87.303830	40.437456
std	13.640433	1.055500e+05	2.572720	7385.292085	402.960219	12.347429
min	17.000000	1.228500e+04	1.000000	0.000000	0.000000	1.000000
25%	28.000000	1.178270e+05	9.000000	0.000000	0.000000	40.000000
50%	37.000000	1.783560e+05	10.000000	0.000000	0.000000	40.000000
75%	48.000000	2.370510e+05	12.000000	0.000000	0.000000	45.000000
max	90.000000	1.484705e+06	16.000000	99999.000000	4356.000000	99.000000

Generated data sample:

	age	workclass	fnlwgt	education	education-num	marital-status	occupation	relationship	race	sex	capital-gain	capital-loss	hours-per-week	native-country	income
0	39	State-gov	77516	Bachelors	13	Never-married	Adm-clerical	Not-in-family	White	Male	2174	0	40	United-States	>50K
1	50	Self-emp-not-inc	83314	Bachelors	13	Married-civ-spouse	Exec-managerial	Husband	White	Male	0	0	13	United-States	>50K
2	38	Private	215646	HS-grad	9	Divorced	Handlers-cleaners	Not-in-family	White	Male	0	0	40	United-States	>50K
3	53	Private	234721	11th	7	Married-civ-spouse	Transportation-machine-op	Husband	Black	Male	0	0	40	United-States	>50K
4	28	Private	338409	Bachelors	13	Married-civ-spouse	Spec-proc	Wife	Black	Female	0	0	40	Cuba	<=50K

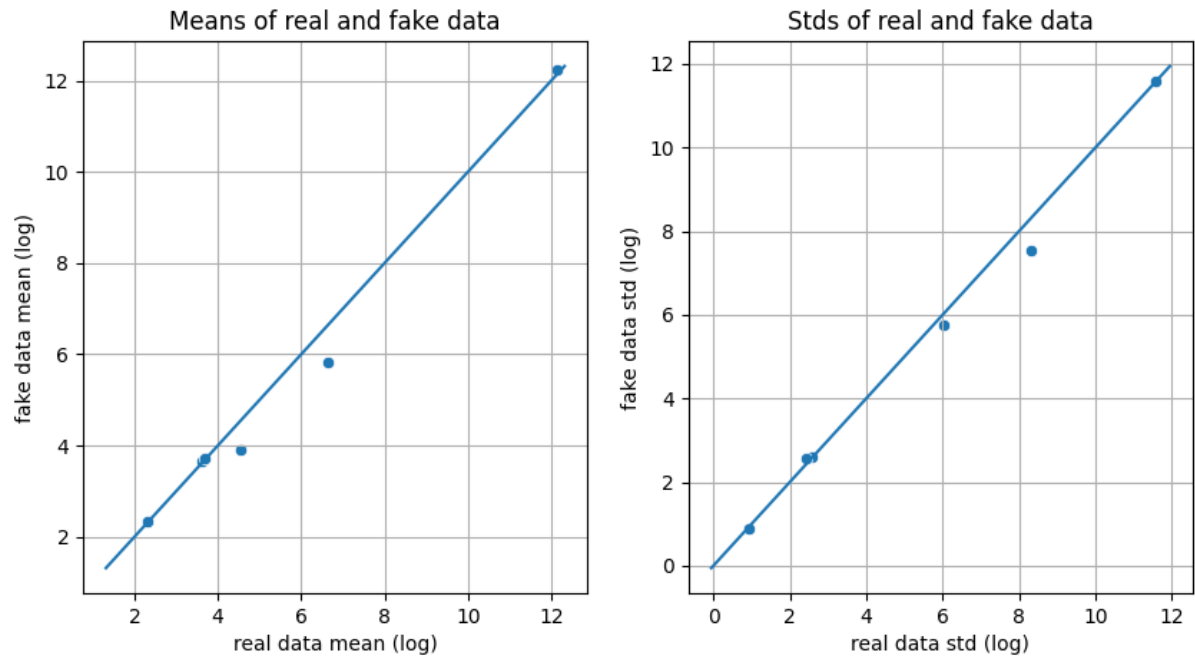
Generated data stats:

	age	fnlwgt	education-num	capital-gain	capital-loss	hours-per-week
count	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000
mean	38.564000	208552.602000	10.463000	337.173000	49.971000	41.321000
std	13.457749	106638.487642	2.420845	1869.919831	318.845814	12.937517
min	15.000000	11232.000000	3.000000	-137.000000	-7.000000	4.000000
25%	29.000000	144385.000000	9.000000	-51.000000	0.000000	40.000000
50%	37.000000	197653.000000	10.000000	-12.000000	3.000000	40.000000
75%	46.000000	249887.000000	13.000000	37.000000	5.000000	41.000000
max	91.000000	788464.000000	17.000000	16270.000000	3280.000000	113.000000

Metrics of real and generated synthetic data

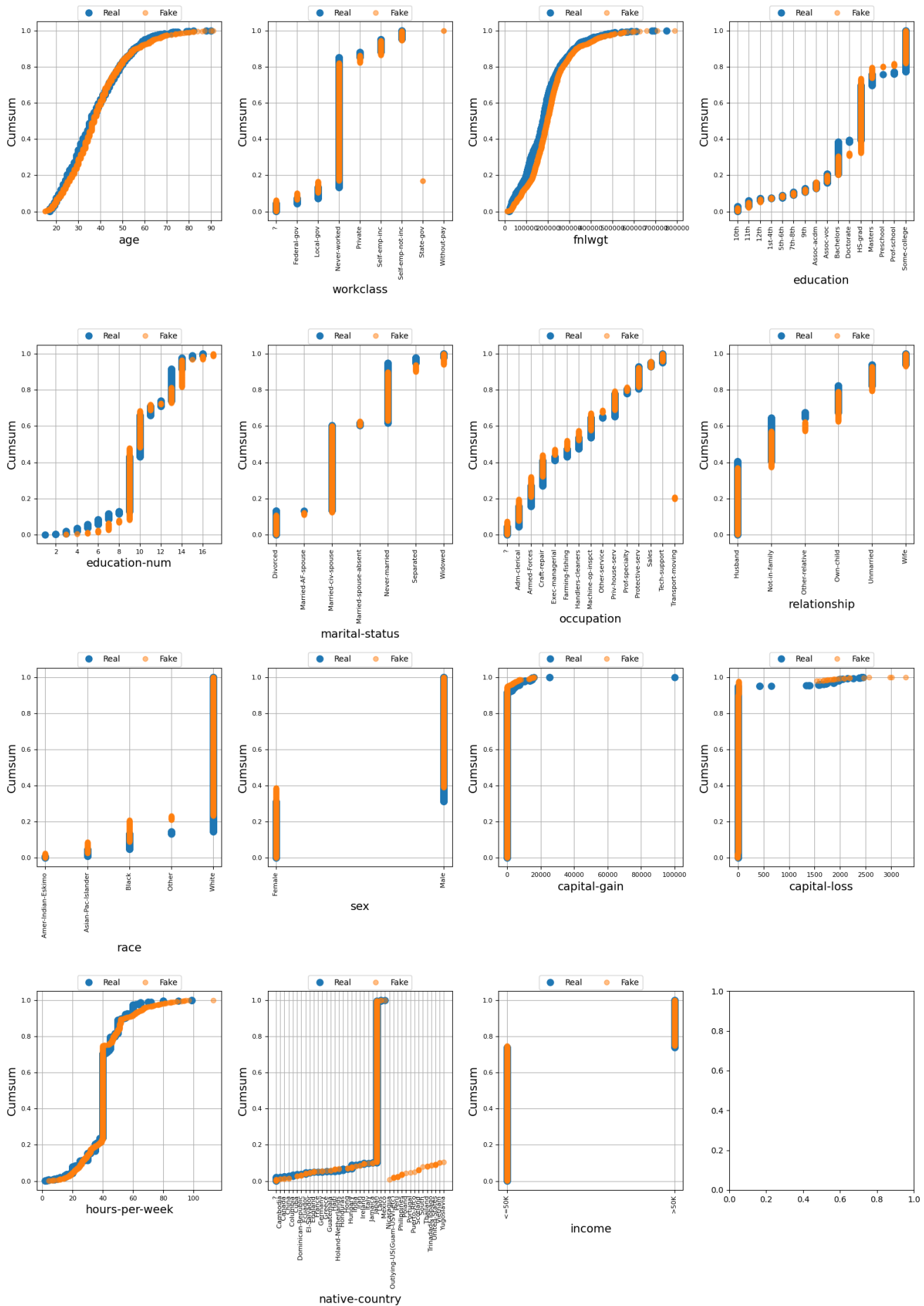
Mean distribution:

Absolute Log Mean and STDs of numeric data



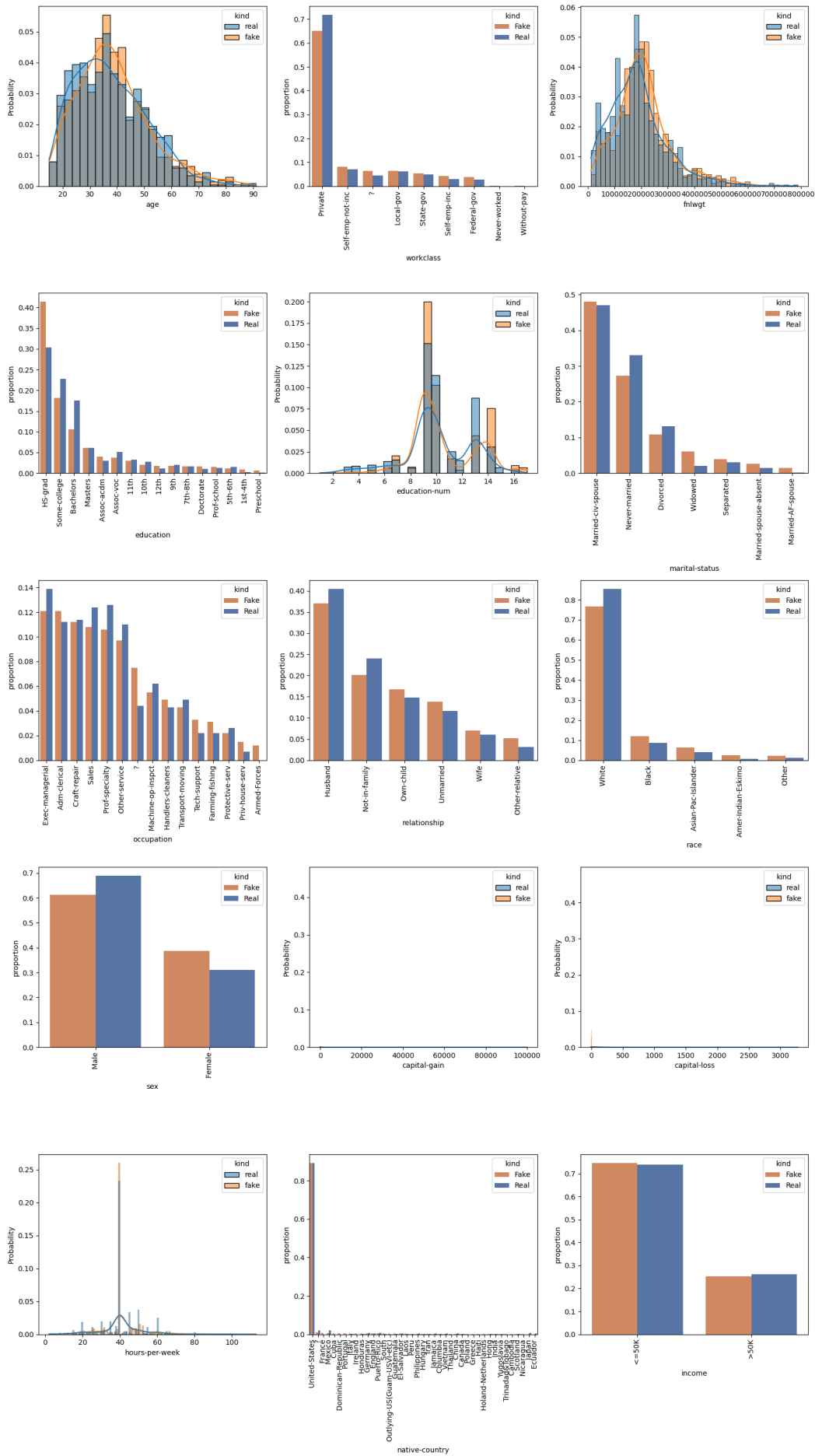
cumulative sums per feature:

Cumulative Sums per feature

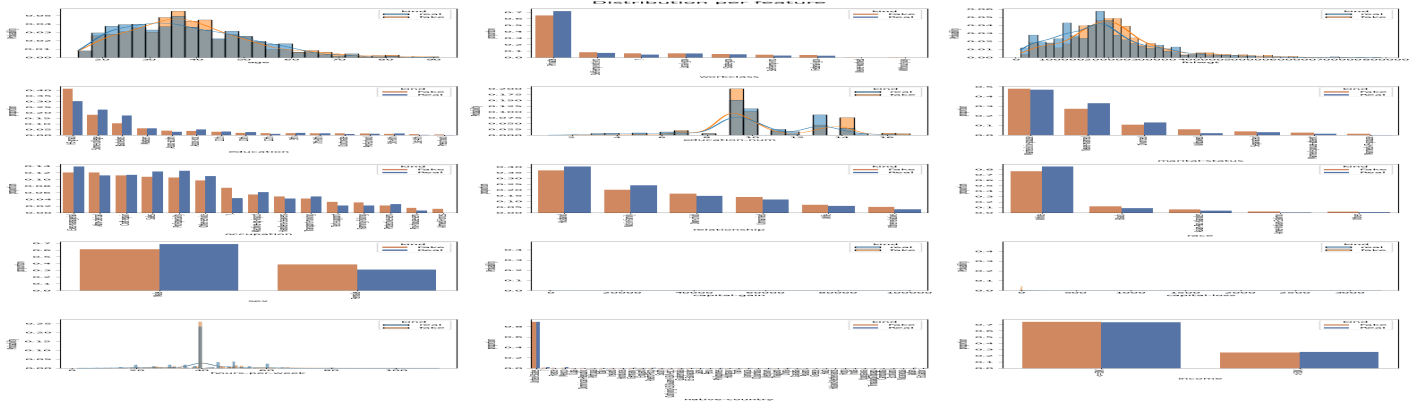


distributions per feature:

Distribution per feature



Correlation difference:



Components of PCA:

