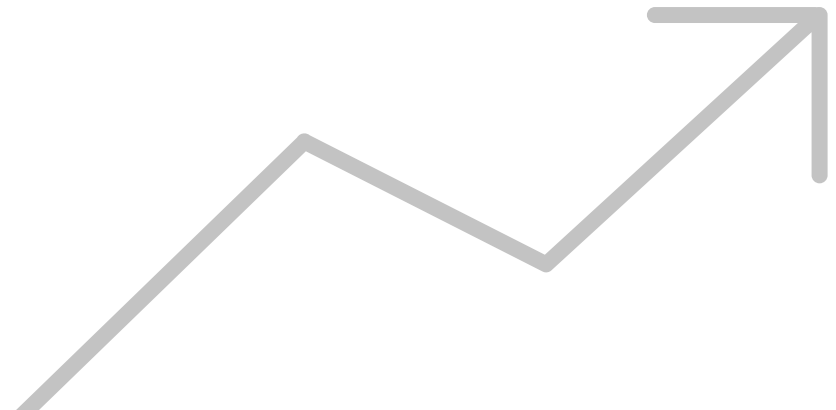


# **HLG-MOS Synthetic Data Challenge**

Team DESTATIS



# Synthetic Data Challenge – Team DESTATIS



Reinhard Tent



Felix Geyer



Michel Reiffert



Hariolf Merkle




Steffen Moritz

# Evaluated Synthesis Methods

Fully Conditional Specification ( <b>FCS</b> )
Generative Adversarial Network ( <b>GAN</b> )
Probabilistic Graphical Models ( <b>Minutemen</b> DP-pgm)
Information Preserving Statistical Obfuscation ( <b>IPSO</b> )
Multivariate non-normal distribution ( <b>Simulated</b> )

# Evaluation Utility Metrics

Method	pMSE	S_pMSE
FCS	0.0497517	1.445132
<b>GAN</b>	0.1186995	3.03237 
Simulated Data	0.2222528	5.111227
IPSO	0.2278401	5.735209
Minuteman	0.2360304	6.078917

# Evaluation Privacy Metrics

Method	Replication Uniques	Disclosure in 1 Var	Disclosure in 2 Var
Simulated Data	0 (0%)	0 (0%)	0 (0%)
Minuteman	2 (0.2%)	0 (0%)	0 (0%)
<b>GAN</b>	10 (1%)	2 (0.2%)	1 (0.1%)
FCS	93 (9.3%)	40 (4%)	7 (0.7%)
IPSO	551 (55.1%)	159 (15.9%)	16 (16%)

# Our Conclusion

SAT DATA	Releasing_Public	Testing_Analysis	Education	Testing_Technology
FCS	NO (privacy)	YES	NO	MAYBE
Simulated Data	NO (utility)	NO	YES	YES
GAN	NO (utility)	NO	YES	YES
IPSO	NO (utility + privacy)	MAYBE	NO	NO
Minutemen	NO (utility)	NO	YES	MAYBE

ACS DATA	Releasing_Public	Testing_Analysis	Education	Testing_Technology
FCS	NO (privacy)	YES	NO	MAYBE
Simulated Data	NO (utility)	NO	YES	YES
GAN	NO (utility)	NO	YES	YES
IPSO	NO (utility + privacy)	MAYBE	NO	NO
Minutemen	NO (utility)	NO	NO	YES 