

MargCTGAN: A “Marginally” Better CTGAN for the Low Sample Regime



Tejumade Afonja, Dingfan Chen, Mario Fritz

CISPA Helmholtz Center for Information Security

Motivation

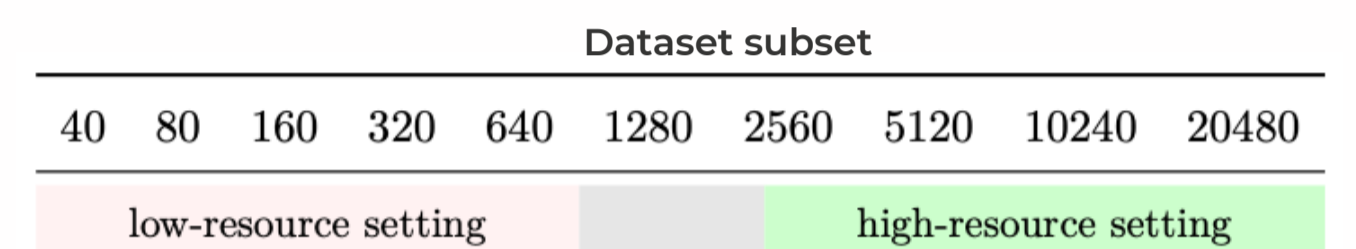
- **Limited exploration** of tabular data generators for **small sample training**
- **Evaluation metrics** for existing generators are **insufficient**
- Synthetic tabular data is beneficial: data augmentation, **sharing**, privacy

Contributions

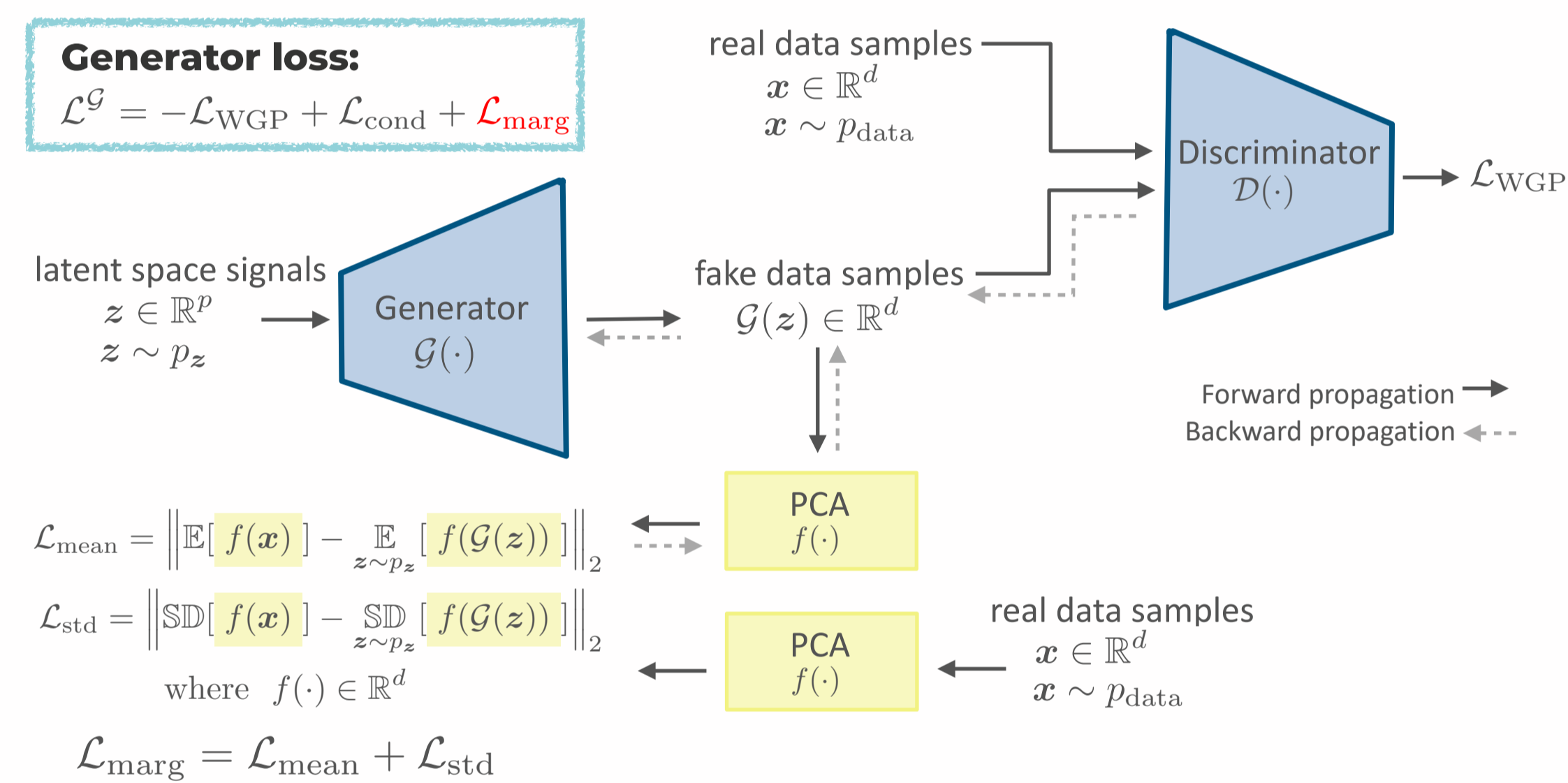
- **Comprehensive evaluation** of generators across key data dimensions — marginal, column-pair, joint and utility — from high to **low sample regime**
- **Enhance CTGAN** model with **moment-matching loss** in de-correlated feature space to capture data distribution characteristics in low sample regime

Datasets and Models

- Datasets: Adult, Census, Texas, and News
- Evaluated Models: CTGAN [1], TVAE [1], TableGAN [2]



MargCTGAN — Moment-matching in De-correlated Feature Space

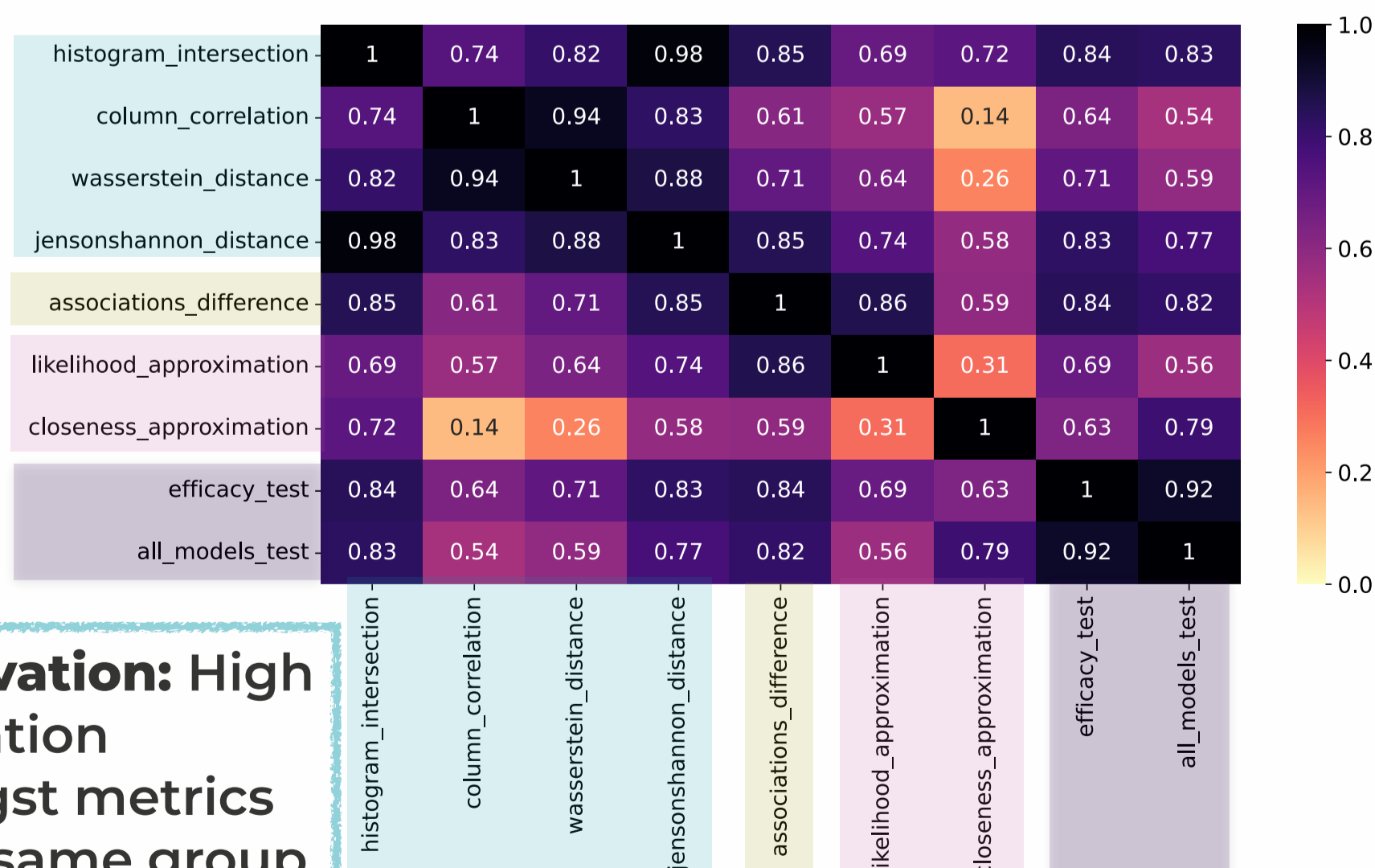


Target: Improve the synthetic data marginal distribution matching while preserving utility, joint and column-pair correlation

Multi-dimensional Evaluation Metrics

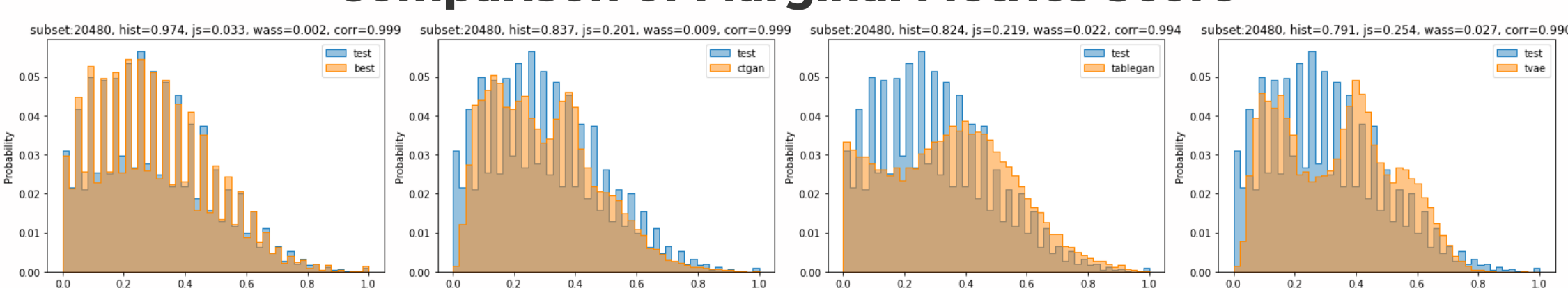
Utility Marginal Column-pair Joint

Correlation Among Metrics

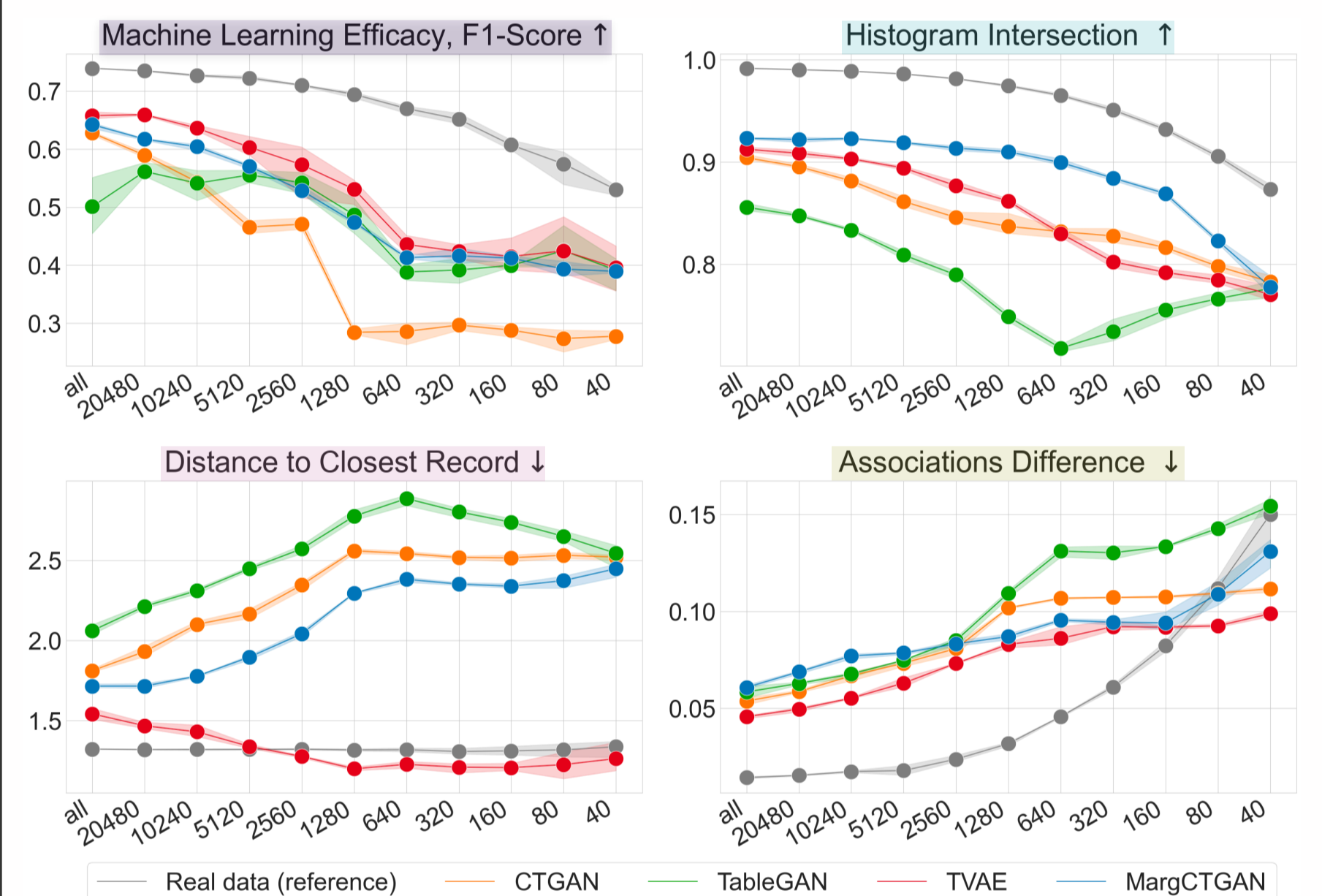


Observation: High correlation amongst metrics in the same group

Comparison of Marginal Metrics Score



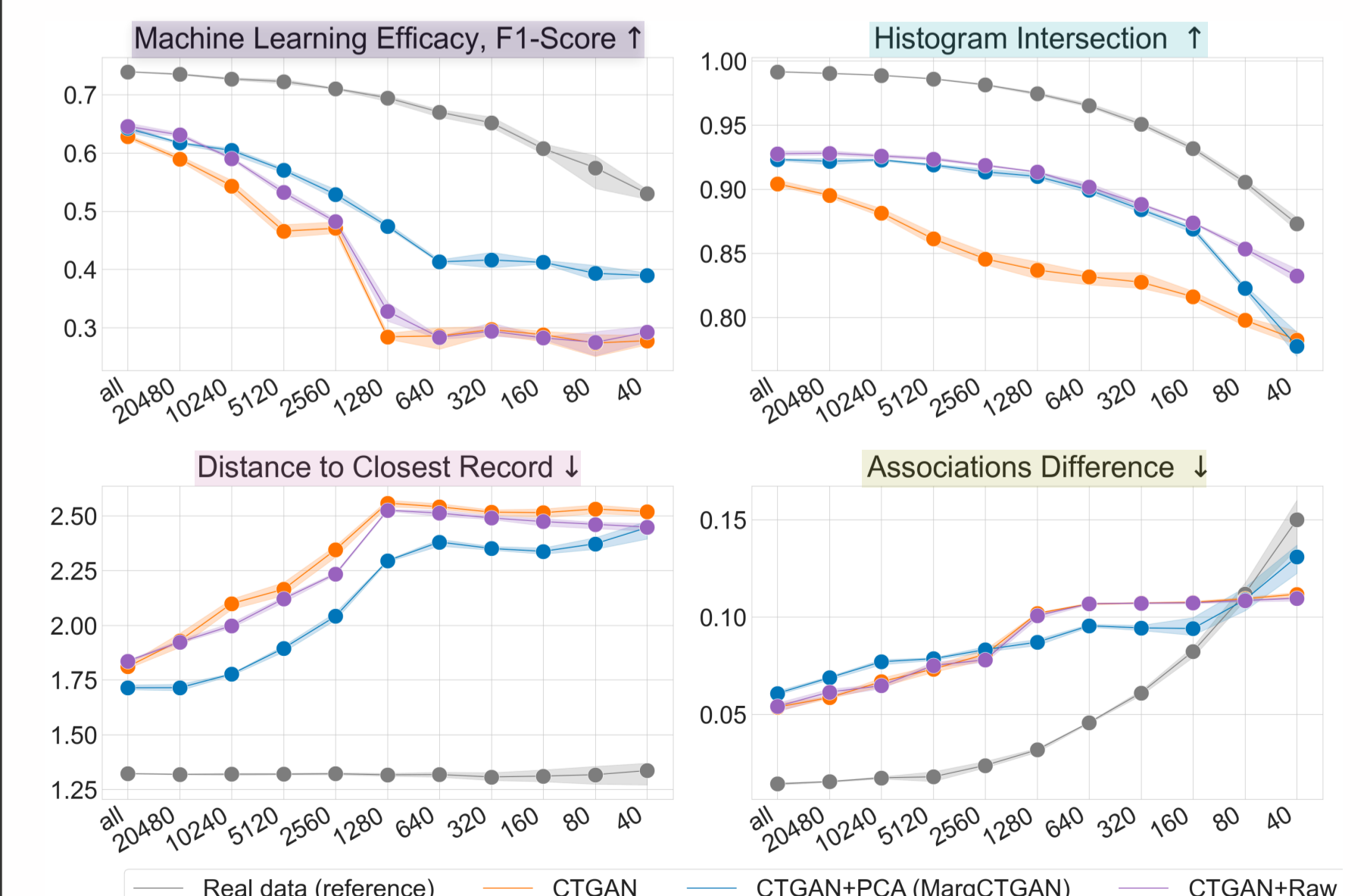
Rethinking Evaluation of Tabular Data Generators



MargCTGAN advantages in low sample regime

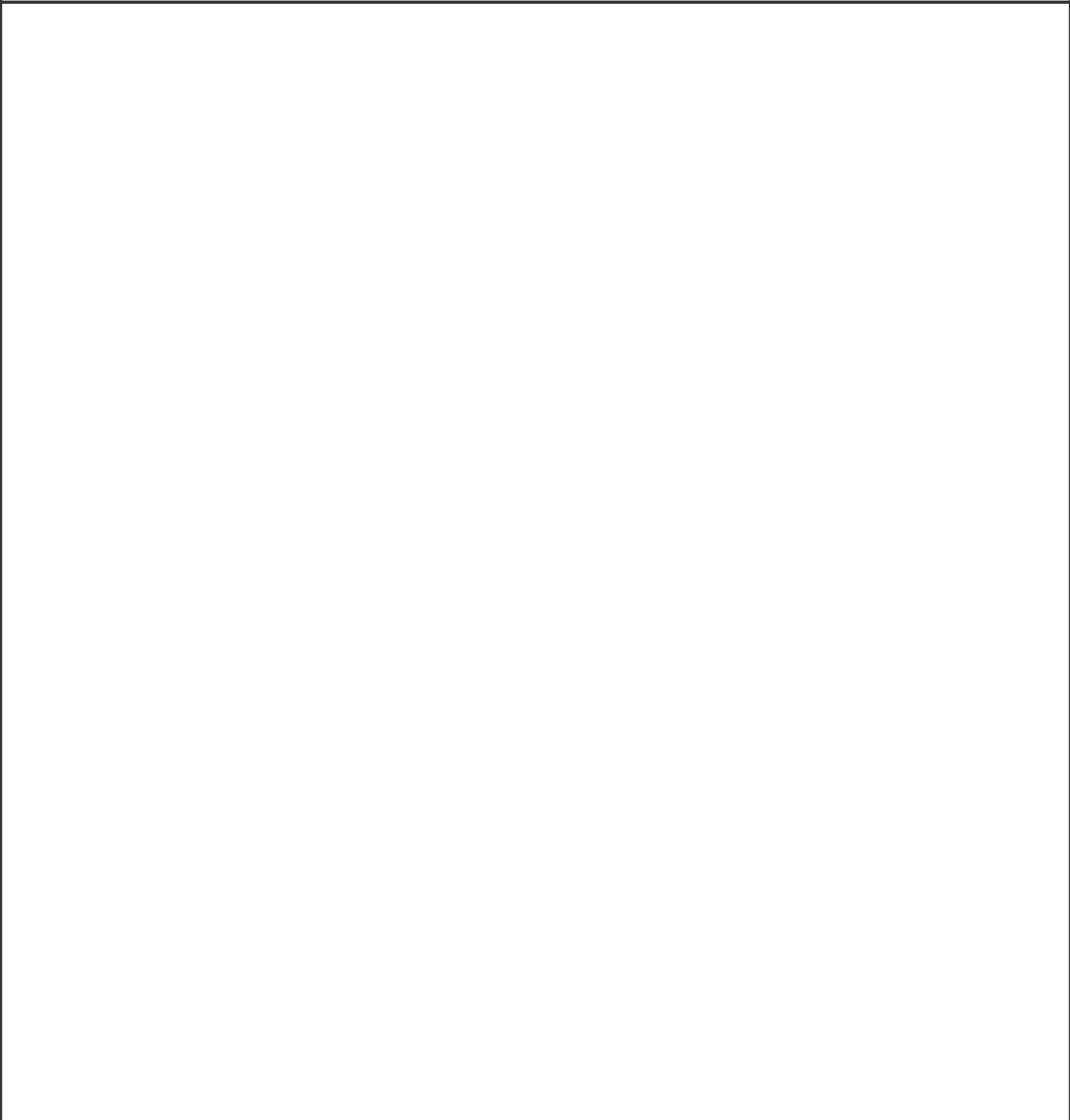
- Better marginal matching
- Competitive downstream utility
- Moderate distance to closest real test
- Small associations difference but sup-optimal from subset 160 downwards

Effect of Moment-matching in PCA Space



References

- Xu, L., Skoularidou, M., Cuesta-Infante, A. and Veeramachaneni, K., 2019. Modeling tabular data using conditional gan. *Advances in Neural Information Processing Systems*, 32
- Park, N., Mohammadi, M., Gorde, K., Jajodia, S., Park, H. and Kim, Y., 2018. Data Synthesis based on Generative Adversarial Networks. *Proceedings of the VLDB Endowment*, 11



HEADLINE OF THE FIRST LINE SINGLE COLUMN

*Here you can insert content in the form of text, image or graphic. You can see how this can look like on the **Example Layout slide**.*

HEADLINE OF THE SECOND LINE SINGLE COLUMN

*Here you can insert content in the form of text, image or graphic. You can see how this can look like on the **Example Layout** slide.*

HEADLINE OF THE THIRD LINE SINGLE COLUMN

*Here you can insert content in the form of text, image or graphic. You can see how this can look like on the **Example Layout** slide.*

**HEADLINE OF THE FIRST LINE:
1. COLUMN**

*Here you can insert content in the form of text, image or graphic. You can see how this can look like on the **Example Layout** slide.*

**HEADLINE OF THE FIRST LINE:
2. COLUMN**

*Here you can insert content in the form of text, image or graphic. You can see how this can look like on the **Example Layout** slide.*

HEADLINE OF THE SECOND LINE: 1. COLUMN

*Here you can insert content in the form of text, image or graphic. You can see how this can look like on the **Example Layout slide**.*

HEADLINE OF THE SECOND LINE: 2. COLUMN

*Here you can insert content in the form of text, image or graphic. You can see how this can look like on the **Example Layout slide**.*

**HEADLINE OF THE THIRD LINE:
1. COLUMN**

*Here you can insert content in the form of text, image or graphic. You can see how this can look like on the **Example Layout** slide.*

**HEADLINE OF THE THIRD LINE:
2. COLUMN**

*Here you can insert content in the form of text, image or graphic. You can see how this can look like on the **Example Layout** slide.*

FIRST LINE: 1. COLUMN	FIRST LINE: 2. COLUMN	FIRST LINE: 3. COLUMN
<p><i>Here you can insert content in the form of text, image or graphic. You can see how this can look like on the Example Layout slide.</i></p>	<p><i>Here you can insert content in the form of text, image or graphic. You can see how this can look like on the Example Layout slide.</i></p>	<p><i>Here you can insert content in the form of text, image or graphic. You can see how this can look like on the Example Layout slide.</i></p>

SECOND LINE: 1. COLUMN	SECOND LINE: 2. COLUMN	SECOND LINE: 3. COLUMN
<p><i>Here you can insert content in the form of text, image or graphic. You can see how this can look like on the Example Layout slide.</i></p>	<p><i>Here you can insert content in the form of text, image or graphic. You can see how this can look like on the Example Layout slide.</i></p>	<p><i>Here you can insert content in the form of text, image or graphic. You can see how this can look like on the Example Layout slide.</i></p>

THIRD LINE: 1. COLUMN	THIRD LINE: 2. COLUMN	THIRD LINE: 3. COLUMN
<p><i>Here you can insert content in the form of text, image or graphic. You can see how this can look like on the Example Layout slide.</i></p>	<p><i>Here you can insert content in the form of text, image or graphic. You can see how this can look like on the Example Layout slide.</i></p>	<p><i>Here you can insert content in the form of text, image or graphic. You can see how this can look like on the Example Layout slide.</i></p>

FIRST LINE: 1. COLUMN	FIRST LINE: 2. COLUMN	FIRST LINE: 3. COLUMN	FIRST LINE: 4. COLUMN
<p><i>Here you can insert content in the form of text, image or graphic. You can see how this can look like on the Example Layout slide.</i></p>	<p><i>Here you can insert content in the form of text, image or graphic. You can see how this can look like on the Example Layout slide.</i></p>	<p><i>Here you can insert content in the form of text, image or graphic. You can see how this can look like on the Example Layout slide.</i></p>	<p><i>Here you can insert content in the form of text, image or graphic. You can see how this can look like on the Example Layout slide.</i></p>

2. LINE: 1. COLUMN	2. LINE: 2. COLUMN	2. LINE: 3. COLUMN	2. LINE: 4. COLUMN
<p><i>Here you can insert content in the form of text, image or graphic. You can see how this can look like on the Example Layout slide.</i></p>	<p><i>Here you can insert content in the form of text, image or graphic. You can see how this can look like on the Example Layout slide.</i></p>	<p><i>Here you can insert content in the form of text, image or graphic. You can see how this can look like on the Example Layout slide.</i></p>	<p><i>Here you can insert content in the form of text, image or graphic. You can see how this can look like on the Example Layout slide.</i></p>

THIRD LINE: 1. COLUMN	THIRD LINE: 2. COLUMN	THIRD LINE: 3. COLUMN	THIRD LINE: 4. COLUMN
<p><i>Here you can insert content in the form of text, image or graphic. You can see how this can look like on the Example Layout slide.</i></p>	<p><i>Here you can insert content in the form of text, image or graphic. You can see how this can look like on the Example Layout slide.</i></p>	<p><i>Here you can insert content in the form of text, image or graphic. You can see how this can look like on the Example Layout slide.</i></p>	<p><i>Here you can insert content in the form of text, image or graphic. You can see how this can look like on the Example Layout slide.</i></p>