PCA & t-SNE: The Essential Visuals

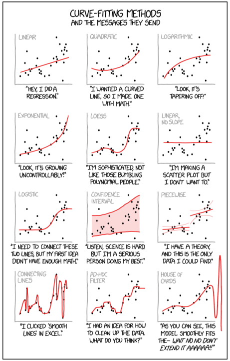
20 Charts Every Data Scientist Should Know

Prof. Dr. Jörg Osterrieder

Methods and Algorithms — MSc Data Science

Spring 2026

The Art of Fitting



XKCD #2048 by Randall Munroe (CC BY-NC 2.5)

Learning Objectives

After this lecture, you will be able to:

- **Interpret** PCA and t-SNE visualizations to assess dimensionality reduction quality
- **Select** the optimal number of components using variance-based diagnostics
- **Compare** linear vs non-linear methods for different data structures

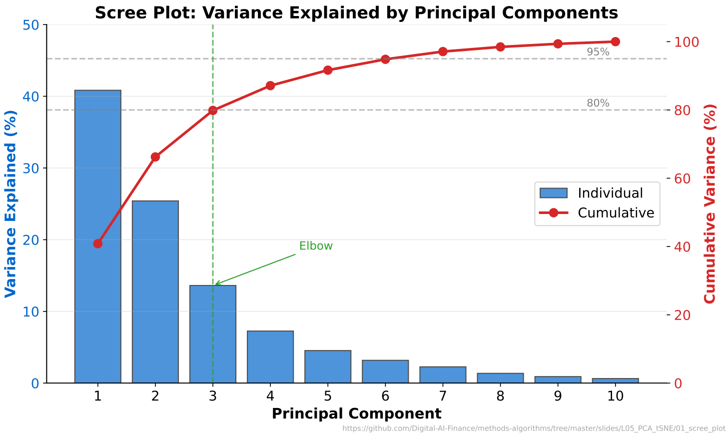
---

We cover 20 essential visualizations spanning PCA foundations, t-SNE mechanics, and advanced topics

Outline

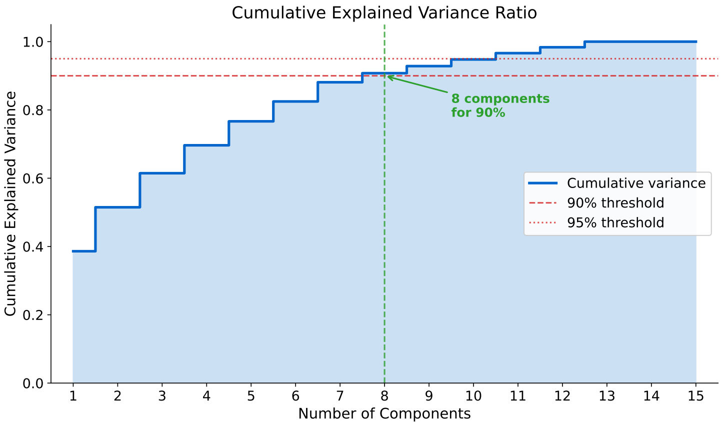
Three parts: PCA Foundations, t-SNE and Non-Linear Methods, Comparison and Advanced Topics

How Much Variance Does Each Component Capture?



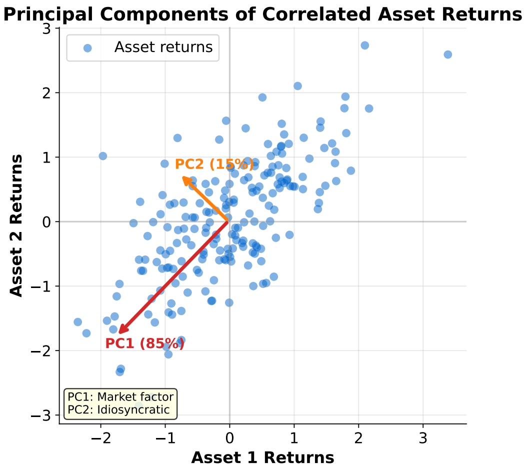
The scree plot shows eigenvalues in descending order—look for the “elbow” where adding components yields diminishing returns

How Many Components Do We Need?



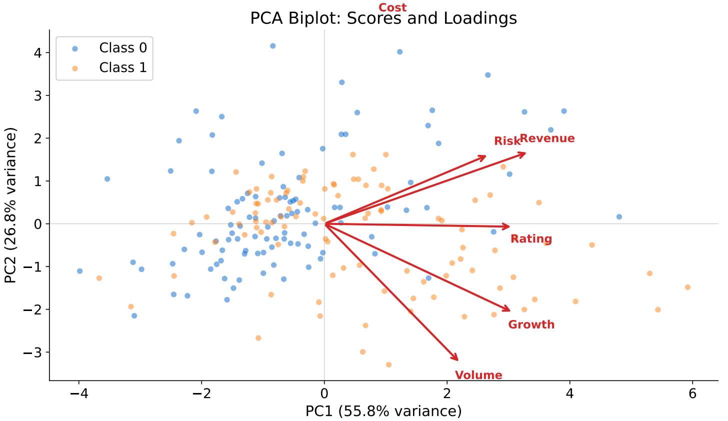
https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L05_PCA_tSNE/top10_09_cumulative_variance

The cumulative variance curve answers the key question: how many PCs are needed for 90% or 95% of total variance?



https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L05_PCA_tSNE/02_principal_components

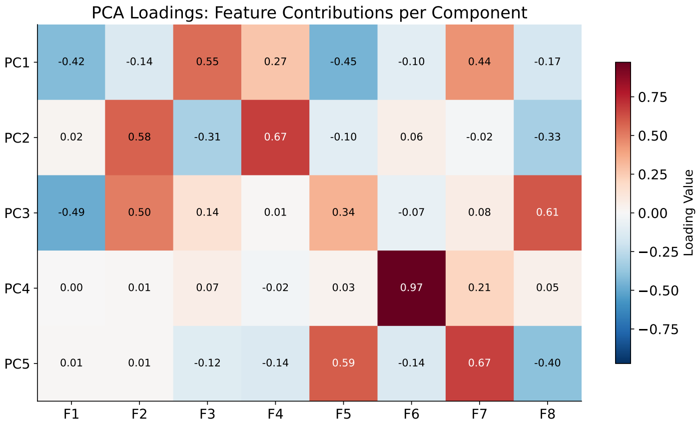
PCA finds orthogonal directions of maximum variance and projects data onto them



https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L05_PCA_tSNE/top10_10_pca_biplot

Biplots overlay data scores with feature loading arrows—arrow direction shows which features align with each PC

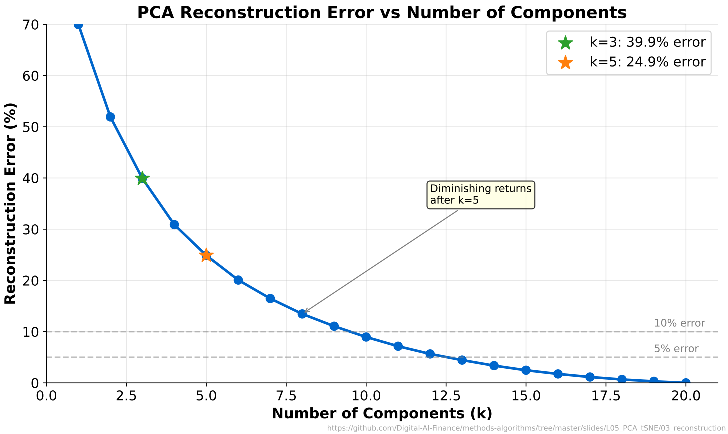
Feature Contributions Across Components



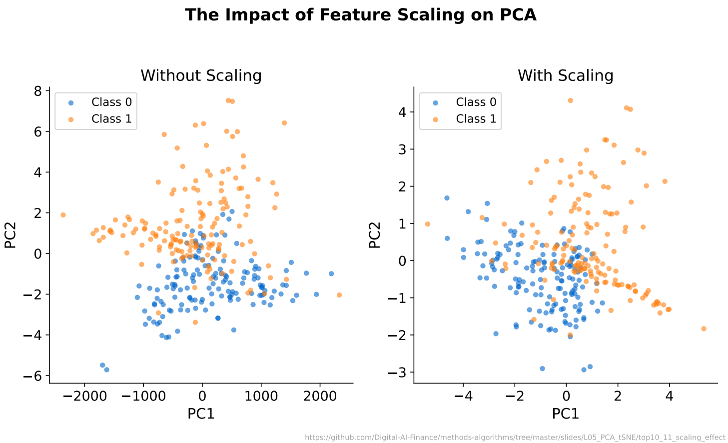
https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L05_PCA_tSNE/top10_14_explained_var_heatmap

Loading heatmaps reveal which features contribute most (positive or negative) to each principal component

Information Loss at Different Dimensions



Reconstruction error quantifies how much information is lost when projecting to fewer dimensions



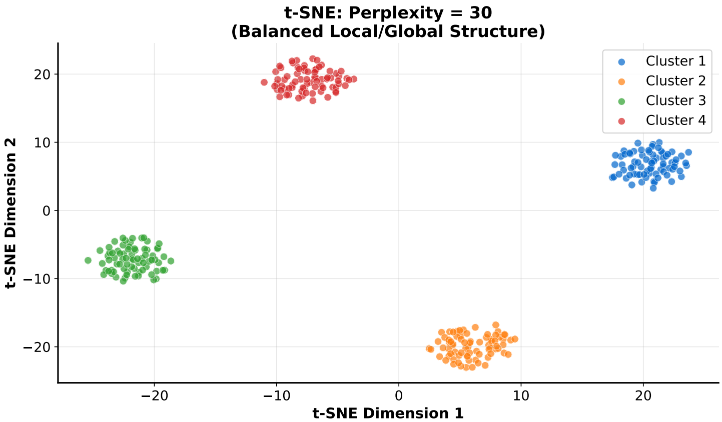
Without scaling, PCA is dominated by high-variance features regardless of their informative content

From Linear to Non-Linear

- PCA captures **global linear structure** but fails on curved manifolds
- t-SNE preserves **local neighbourhood structure** in low dimensions
- Kernel PCA extends PCA to non-linear settings via the kernel trick

When data lies on a non-linear manifold, linear projections distort the true structure

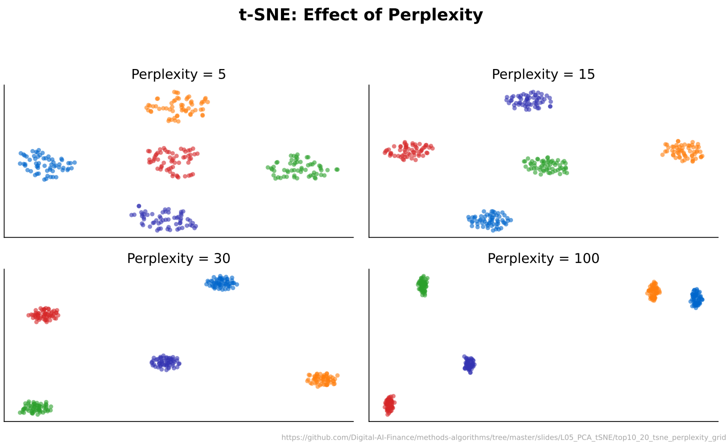
t-SNE: Preserving Local Structure



https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L05_PCA_tSNE/04b_tsne_perplexity_30

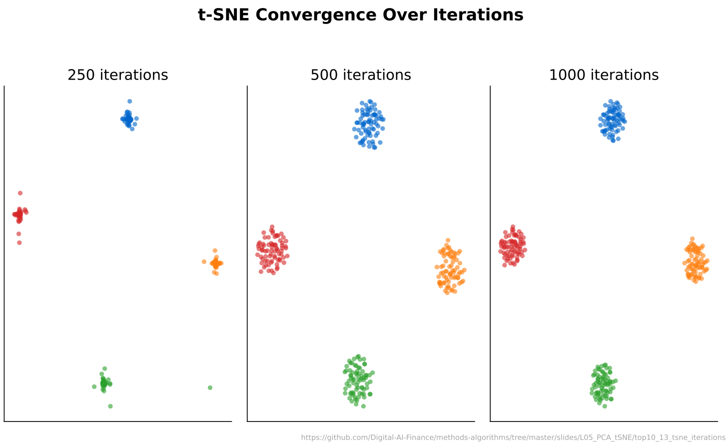
t-SNE converts pairwise similarities to probabilities and minimizes KL divergence between high- and low-dimensional distributions

Effect of Perplexity on t-SNE



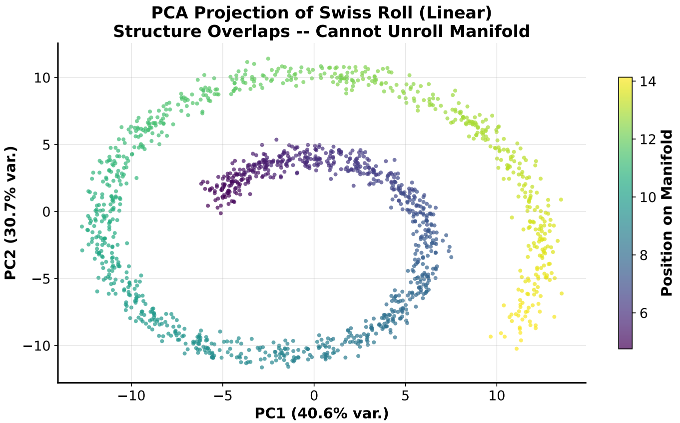
Perplexity controls the effective number of neighbours—low values emphasise local, high values emphasise global structure

t-SNE Convergence Over Iterations



Too few iterations produce noisy embeddings; convergence typically requires 500–1000 iterations

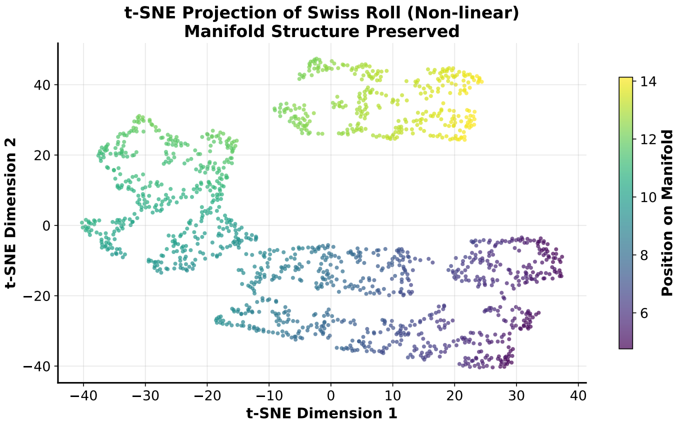
When PCA Fails: The Swiss Roll



https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L05_PCA_tSNE/05a_pca_swiss_roll

PCA projects the Swiss Roll onto a plane, collapsing the manifold structure and mixing distant points

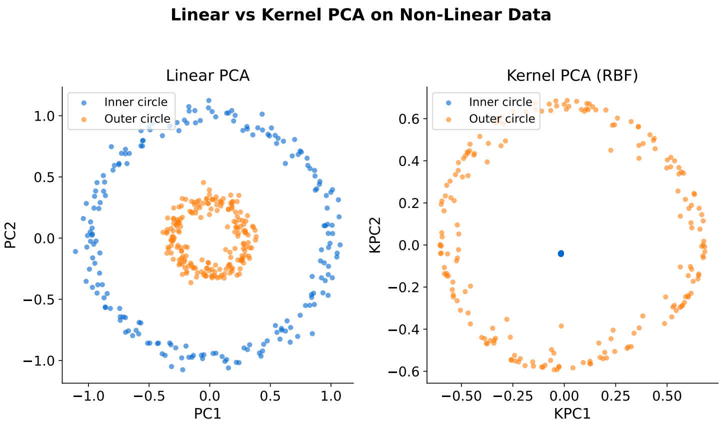
When t-SNE Succeeds: Unrolling the Manifold



https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L05_PCA_tSNE/05b_tsne_swiss_roll

t-SNE preserves local distances along the manifold, successfully "unrolling" the Swiss Roll

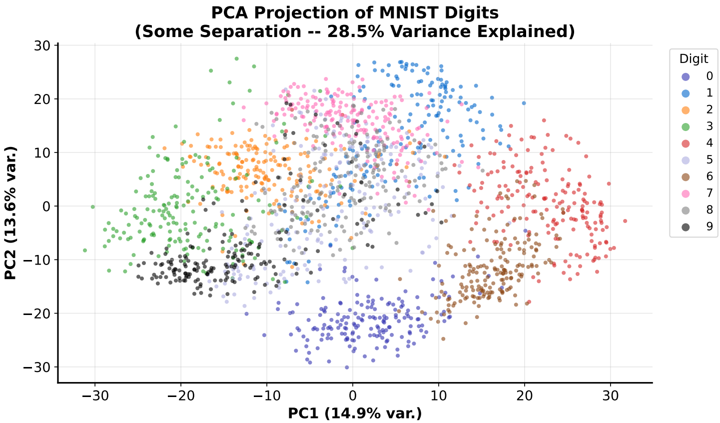
Kernel PCA: Linear Methods on Non-Linear Data



https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L05_PCA_tSNE/top10_12_kernel_pca

Kernel PCA maps data to a higher-dimensional space where linear separation becomes possible

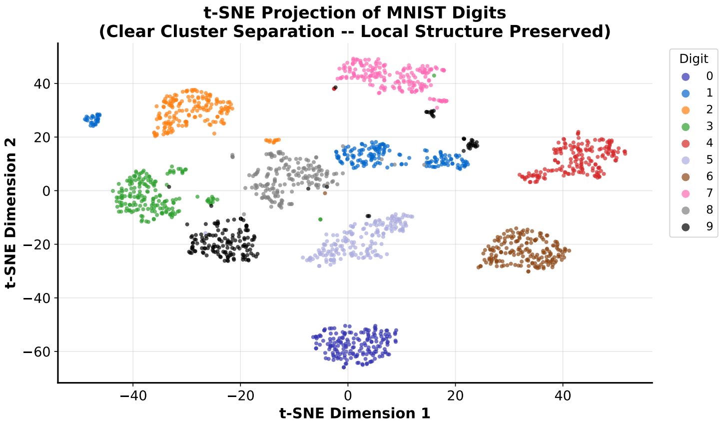
PCA Cluster Projection



https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L05_PCA_tSNE/06b_pca_cluster_projection

PCA preserves global variance but may overlap clusters that are separated in higher dimensions

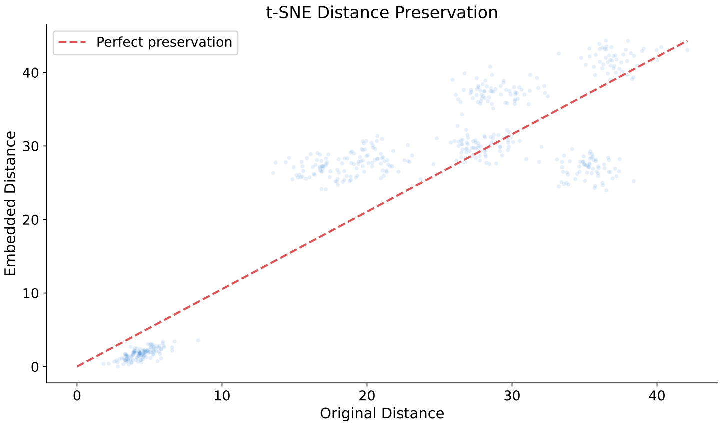
t-SNE Cluster Projection



https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L05_PCA_tSNE/06c_tsne_cluster_projection

t-SNE excels at separating clusters by preserving local neighbourhood relationships

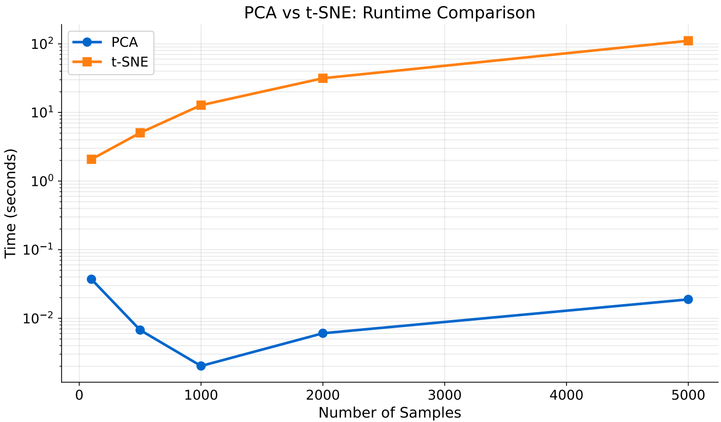
Does t-SNE Preserve Distances?



https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L05_PCA_tSNE/top10_16_tsne_distance_preservation

t-SNE preserves local distances well but distorts global distances—inter-cluster gaps are not meaningful

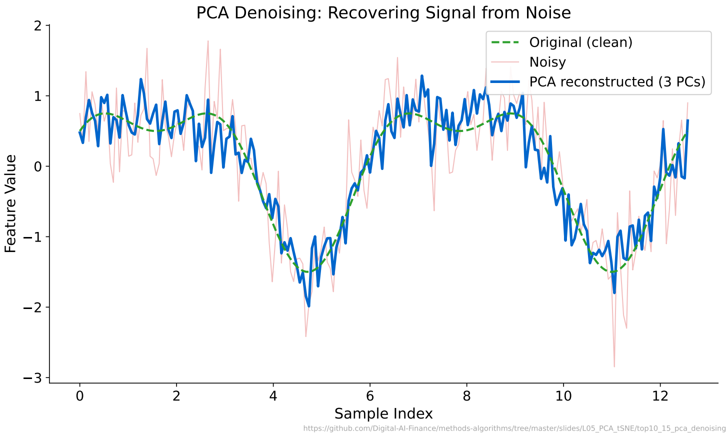
PCA vs t-SNE: Runtime Comparison



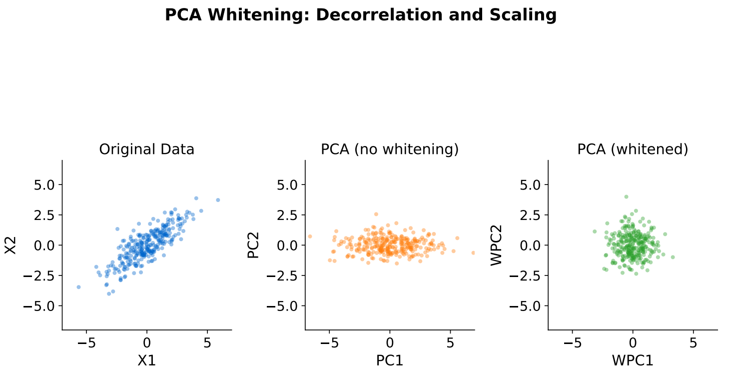
https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L05_PCA_tSNE/top10_17_pca_vs_tsne_runtime

PCA scales linearly; t-SNE scales quadratically. For large datasets, run PCA first to reduce dimensions before t-SNE

PCA for Denoising: Recovering Signal from Noise



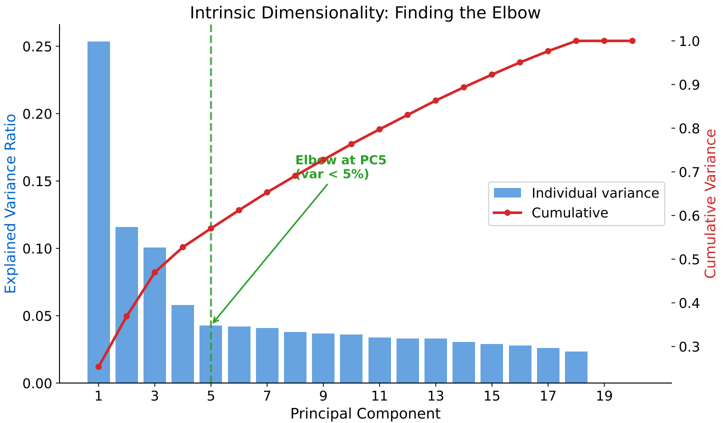
Projecting onto top PCs and reconstructing removes noise captured by lower components



https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L05_PCA_tSNE/top10_19_pca_whitening

Whitening decorrelates features and scales them to unit variance—often used as preprocessing for neural networks

Finding the Intrinsic Dimensionality



https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L05_PCA_tSNE/top10_18_intrinsic_dimensionality

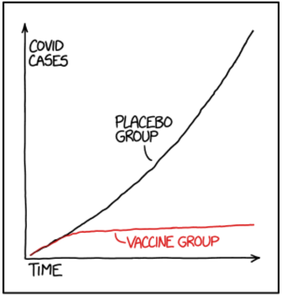
The elbow in explained variance reveals the true number of informative dimensions in the data

20 Charts You Should Know

- ✓ Scree Plot
- ✓ Cumulative Variance
- ✓ 2D Projection
- ✓ Biplot
- ✓ Loading Heatmap
- ✓ Reconstruction
- ✓ Scaling Effect
- ✓ Kernel PCA
- ✓ t-SNE Iterations
- ✓ Perplexity Grid
- ✓ t-SNE Perplexity
- ✓ Swiss Roll (PCA)
- ✓ Swiss Roll (t-SNE)
- ✓ PCA Clusters
- ✓ t-SNE Clusters
- ✓ Distance Preservation
- ✓ Runtime
- ✓ Denoising
- ✓ Whitening
- ✓ Intrinsic Dimensionality

These 20 visualizations cover the foundations, mechanics, and interpretation of dimensionality reduction

The Signal in the Noise



STATISTICS TIP: ALWAYS TRY TO GET DATA THAT'S GOOD ENOUGH THAT YOU DON'T NEED TO DO STATISTICS ON IT