

STAT 847: Reading Assignment 3

DUE: Saturday March 9, 2024 by 11:59pm Eastern

Your assignment must be submitted by the due date listed at the top of this document, and it must be submitted electronically in .pdf format via Crowdmark.

For this reading assignment, read chapter 4, “Nonlinear Dimension Reduction for Visualization”, of Modern Dimension Reduction by Philip D. Waggoner. This can be found in LEARN and in the U Waterloo Library.

Q1 (1 point): What are two methods used to assess the patterns in the data visually?

Q2 (1 point): What’s the difference between t-SNE and the more traditional dimension reduction method of PCA?

Q3 (1 point): What’s an important difference between t-SNE and UMAP?

Q4 (1 point): How is the difference between the full and the reduced data measured when using t-SNE?

Q5 (2 point): Would running a t-SNE twice give exactly the same results twice? Why or why not?

Q6 (1 point): How does increasing the perplexity change the output of the t-SNE?

Q7 (2 points): In UMAP, what are the two searches used for learning the high-dimensional structure of the data?

Q8 (1 point): How are the hyperparameters in t-SNE and in UMAP set to provide an ideal picture of the data.