



The code matches the assignment's mathematical workflows and required outputs. In Part 1 the data are centered, a  $2 \times 2$  covariance is computed, and `np.linalg.eigh` is used with eigenvalues reordered descending; the leading eigenvector aligns with sklearn's PCA components up to sign, and the eigenvalues equal sklearn's explained\_variance. The printed eigenvalue ratios equal explained\_variance\_ratio, and the manual 1D projection  $Xc @ v1$  matches `x_pca[:,0]` up to sign, confirming the derivation. The saved `projection_1D.png` shows the expected ordering/spread along the dominant variance direction, consistent with the original diagonal trend.

In Part 2 `PCA(n_components=2)` on the  $6000 \times 784$  MNIST matrix produces a  $6000 \times 2$  embedding that clusters by digit in `MNIST_PCA_2D.png`. Reconstruction uses `inverse_transform`, yielding `MNIST_reconstructed_2PC.png` that is a blurred but faithful low-rank approximation of `MNIST_original.png`, and selecting coordinates in PCA space generates `MNIST_reconstructed_1_from_coord.png` as expected. In Part 3 NMF (`nndsvda` init, non-negative X) returns W,H; rows of W are normalized to sum to 1, then averaged by clade. The TSV is alphabetically sorted by clade, ancestry columns are ordered by prevalence across clades, values are rounded to two decimals, and the file is saved as `dogs_ancestry_summary.tsv`. Shapes, filenames, and prints match the specification, indicating correctness.