Progress Report

Kuei-Yueh (Clint) Ko

Duke Data+

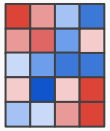
Design Tutorial: Scanpy and Scikit Learn

Genes / Features

Cells (observation)

Matrix of Counts

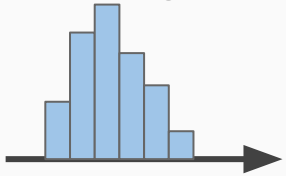
Visualization: Heatmap



Sklearn.preprocessing

- Standardization
- Scaling to (0, 1)

Visualization: Histogram



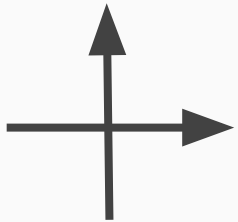
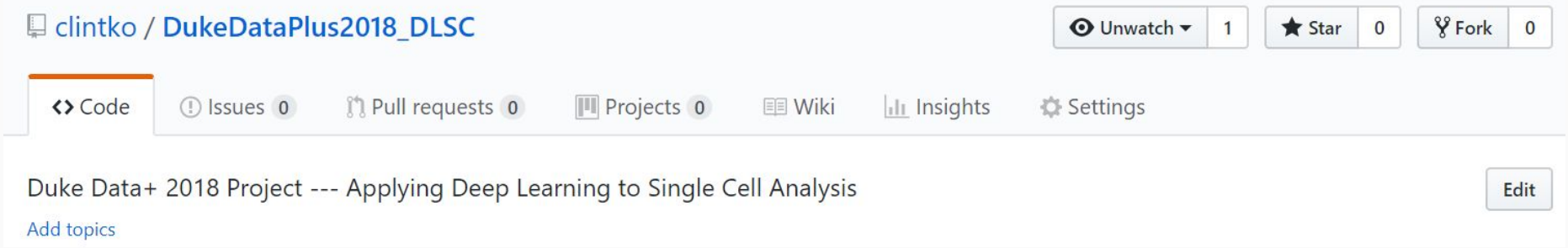
Dimensional Reduction

- sklearn.decomposition.PCA
- sklearn.manifold.TSNE

Hierarchical clustering

- Sklearn.cluster. AgglomerativeClustering

Visualization: scatterplot



Create Duke Data+ Github Repository