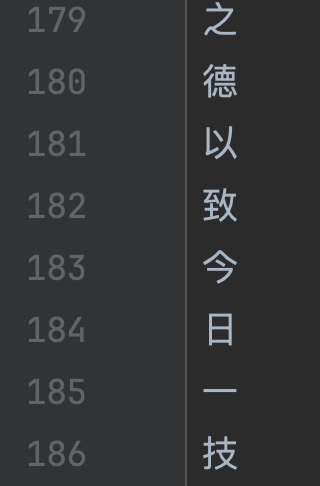
红楼梦作者分析--期中进展

计算机2102 刘钰 2111640206

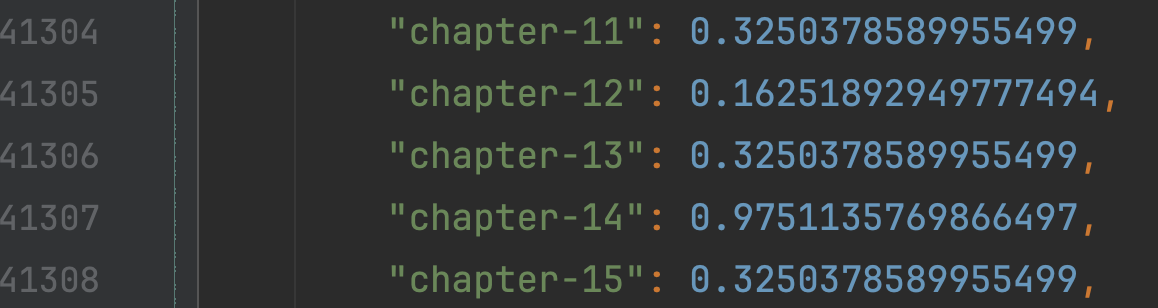
一、 按字

a. 1-gram

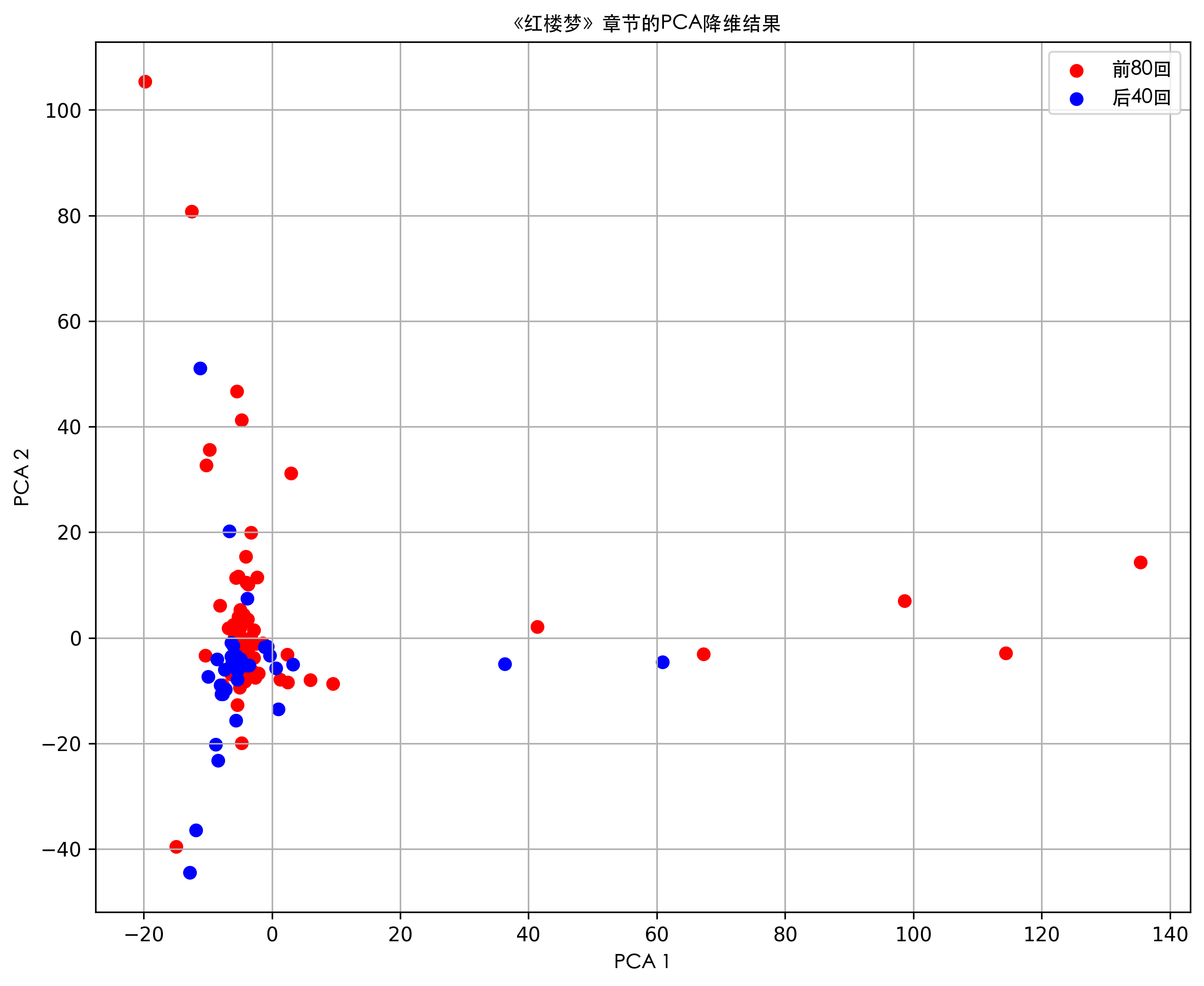
1. 数据预处理



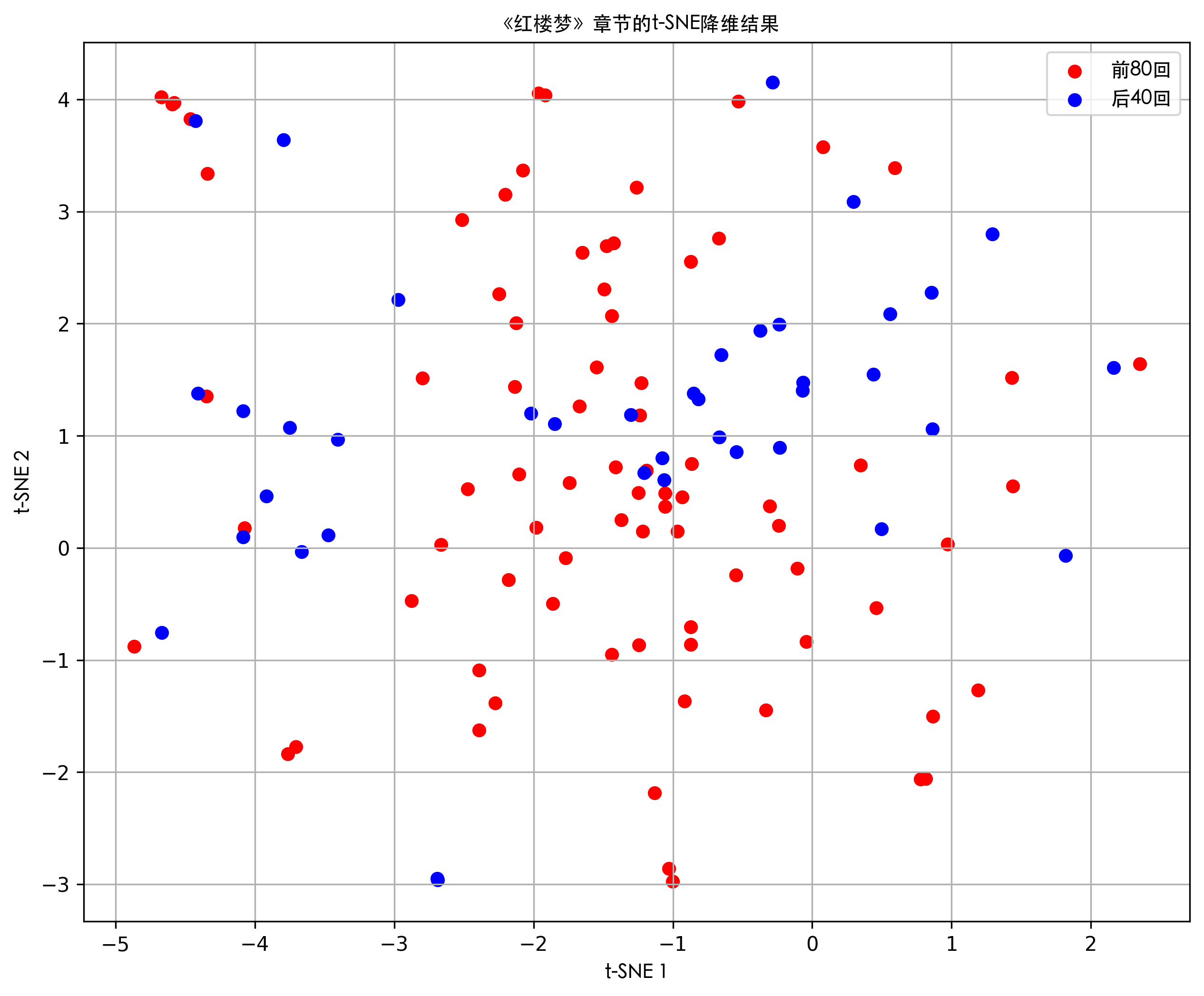
2. tfidf



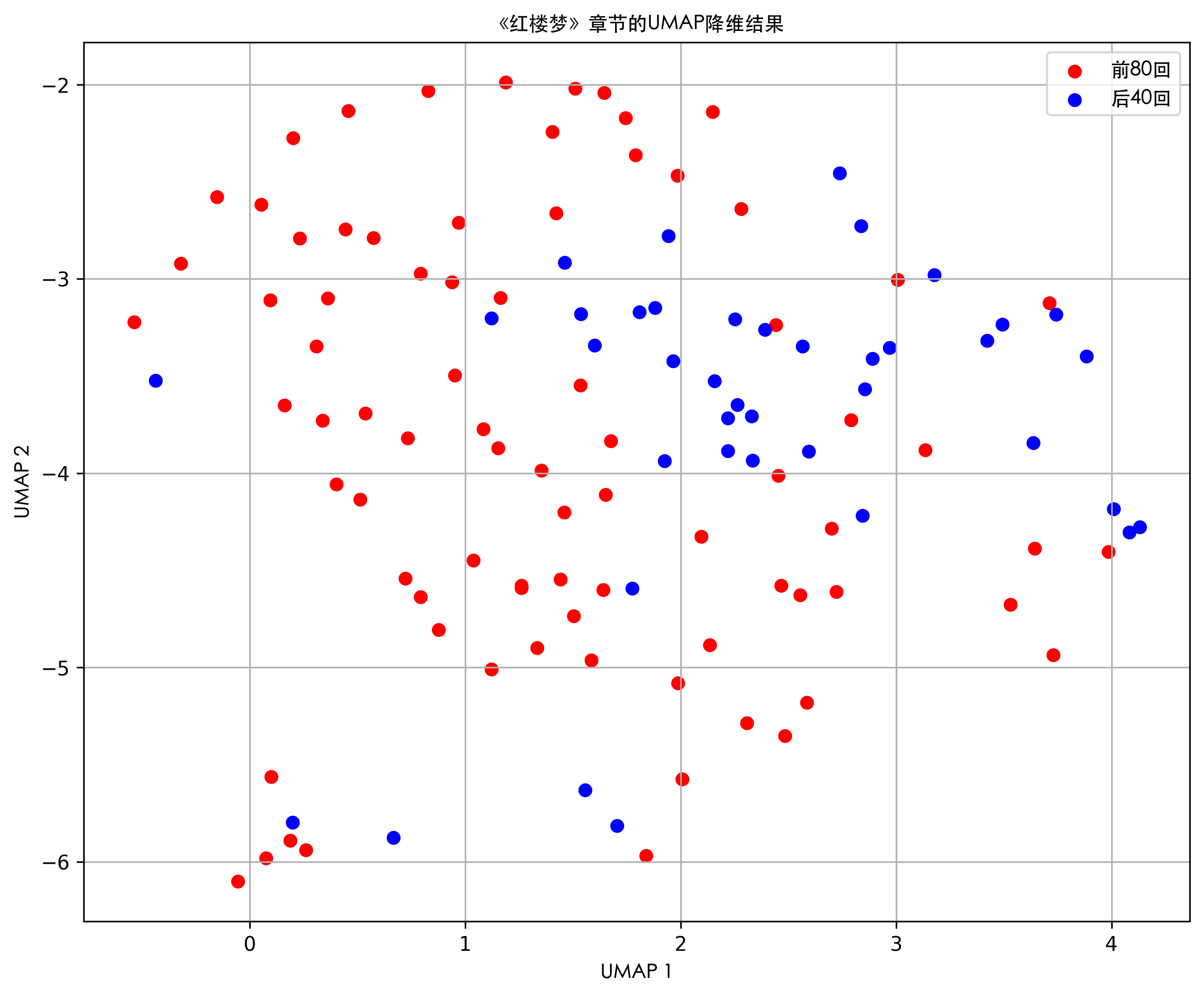
3. pca



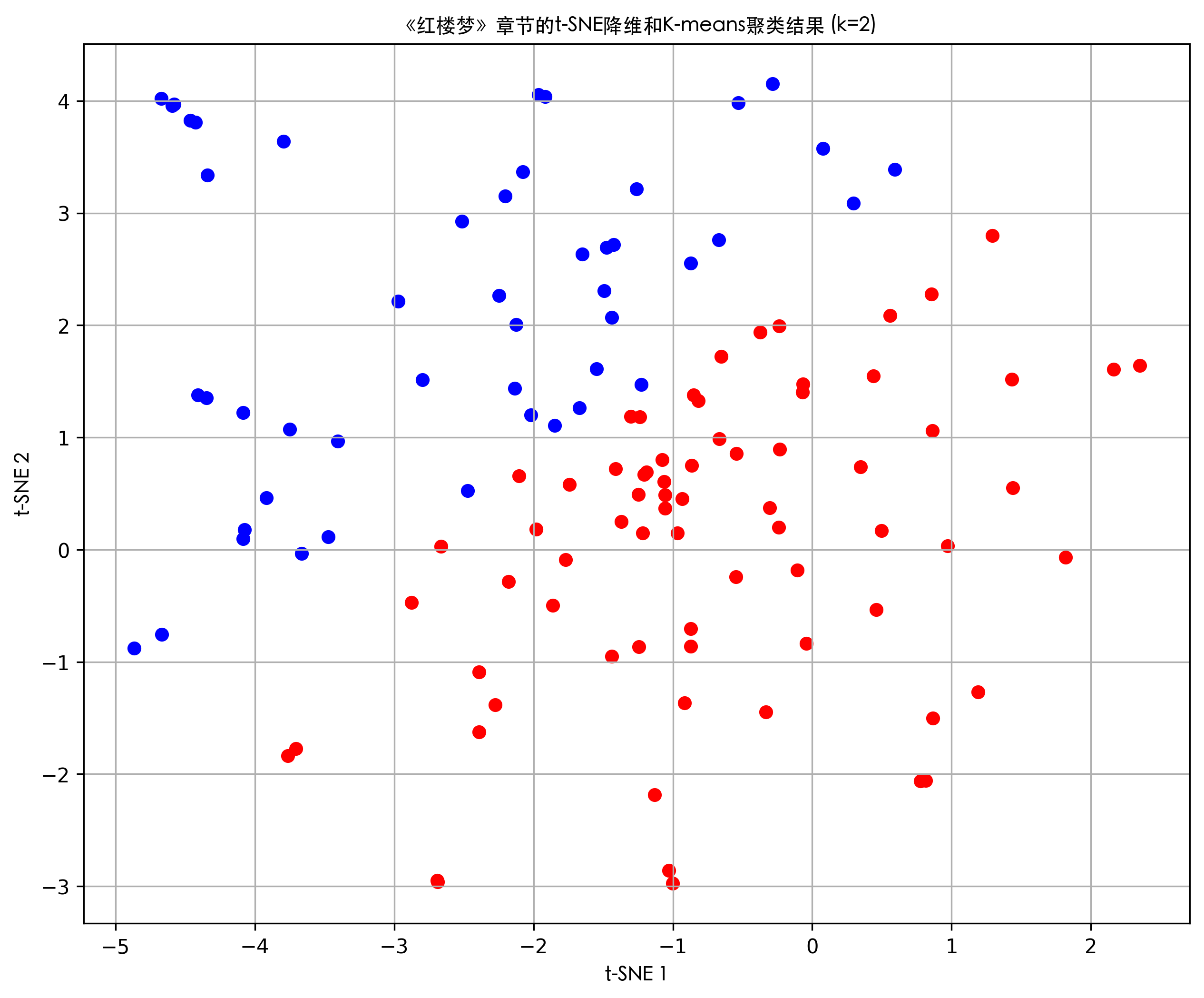
4. tsne



5. umap



6. kmeans

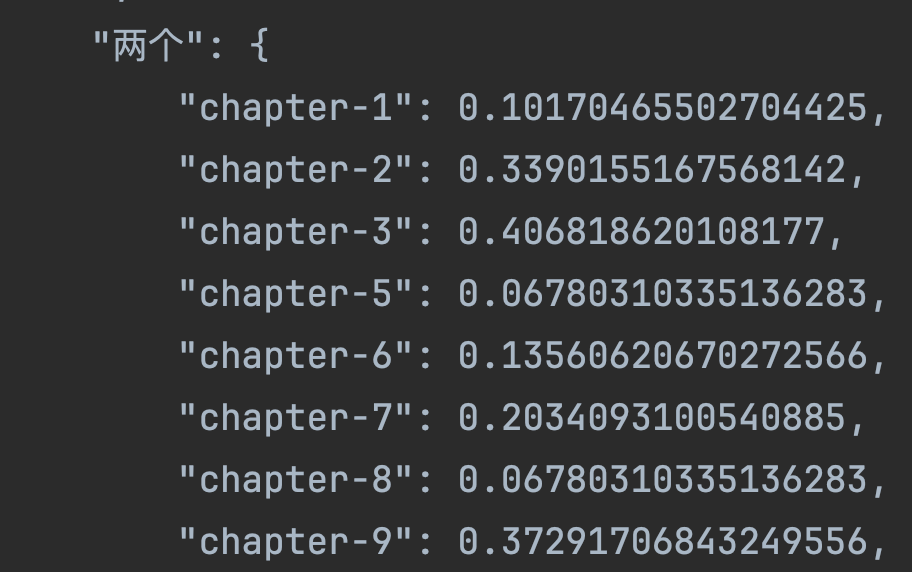


b. 2-gram

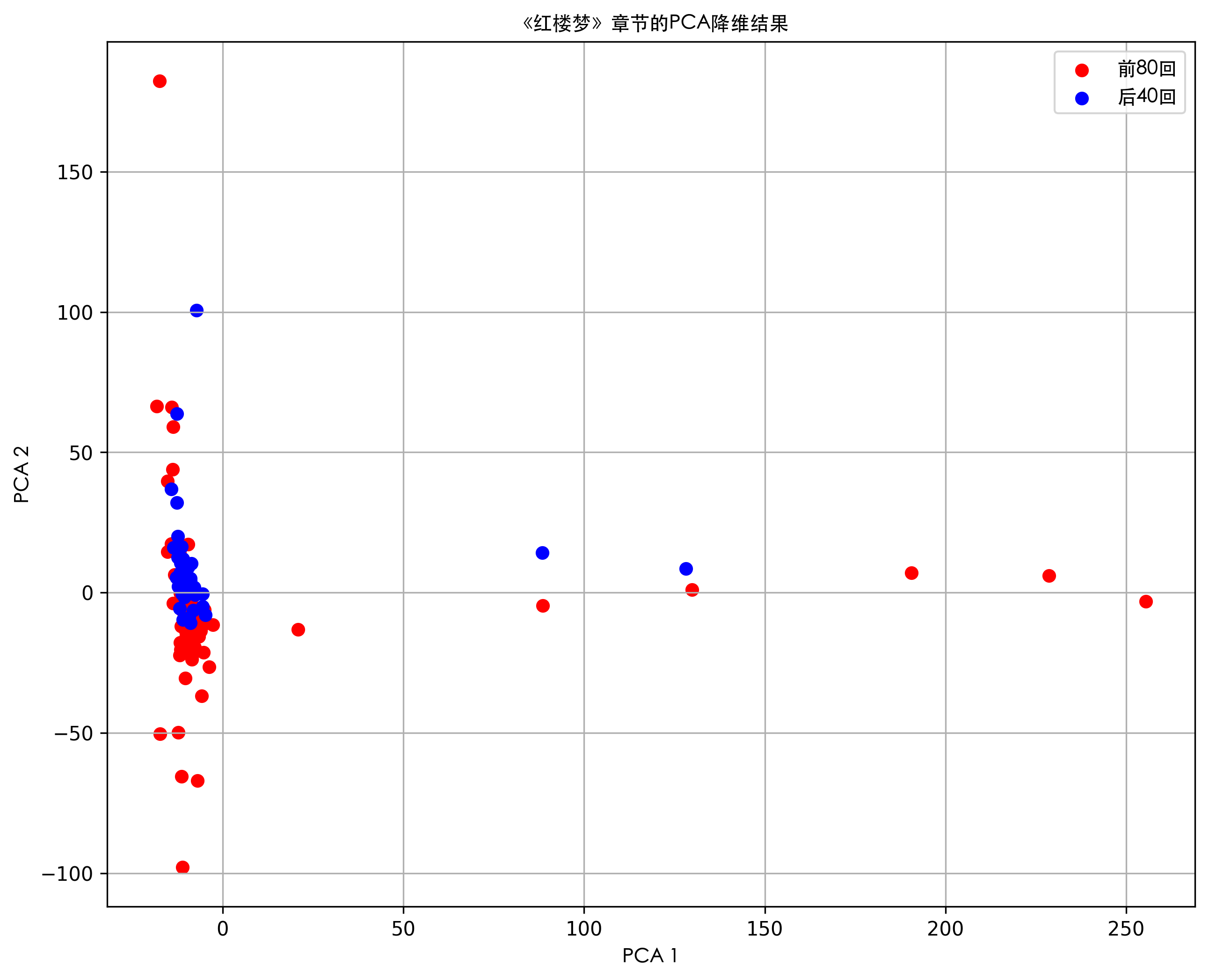
1. 数据预处理



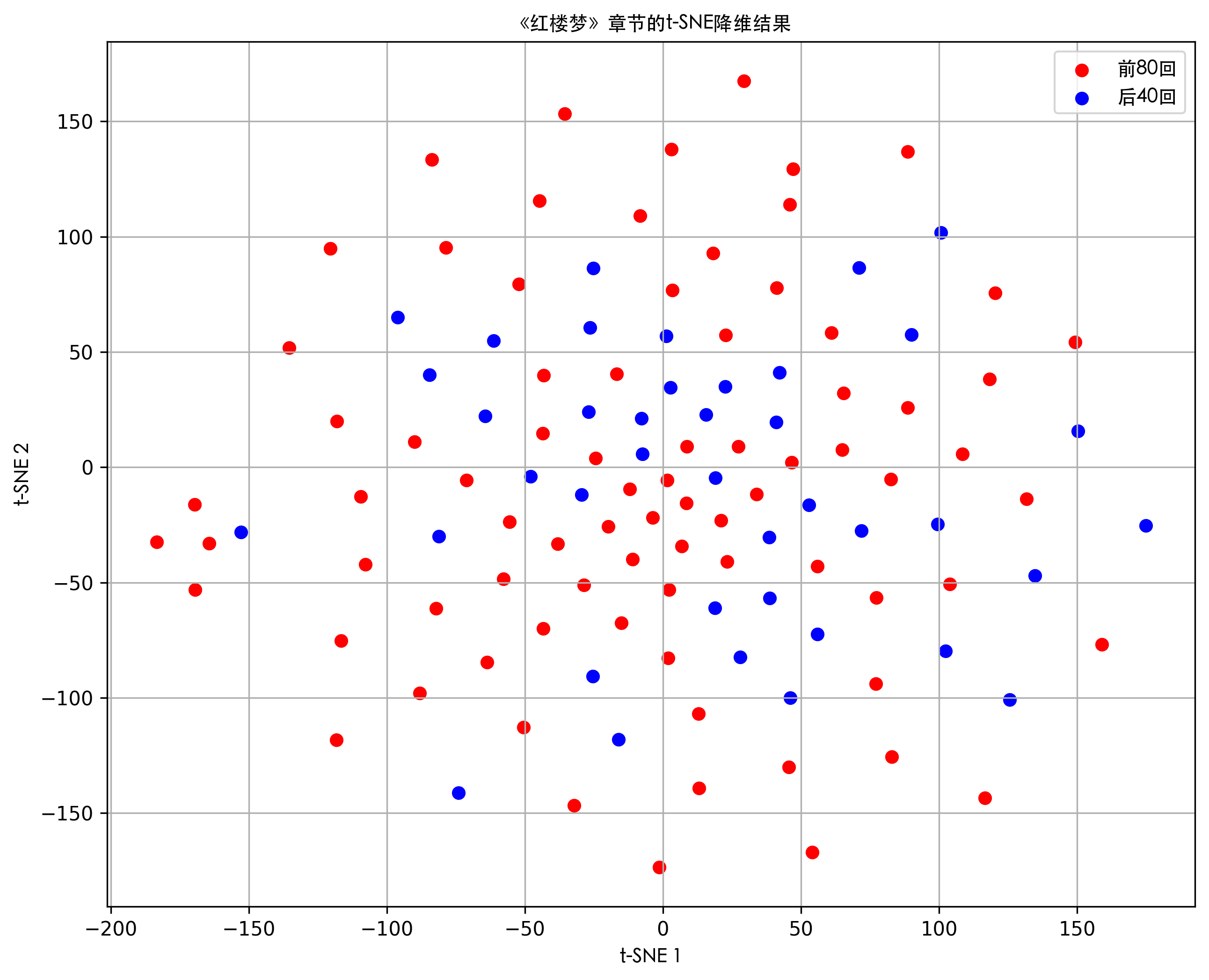
2. tfidf



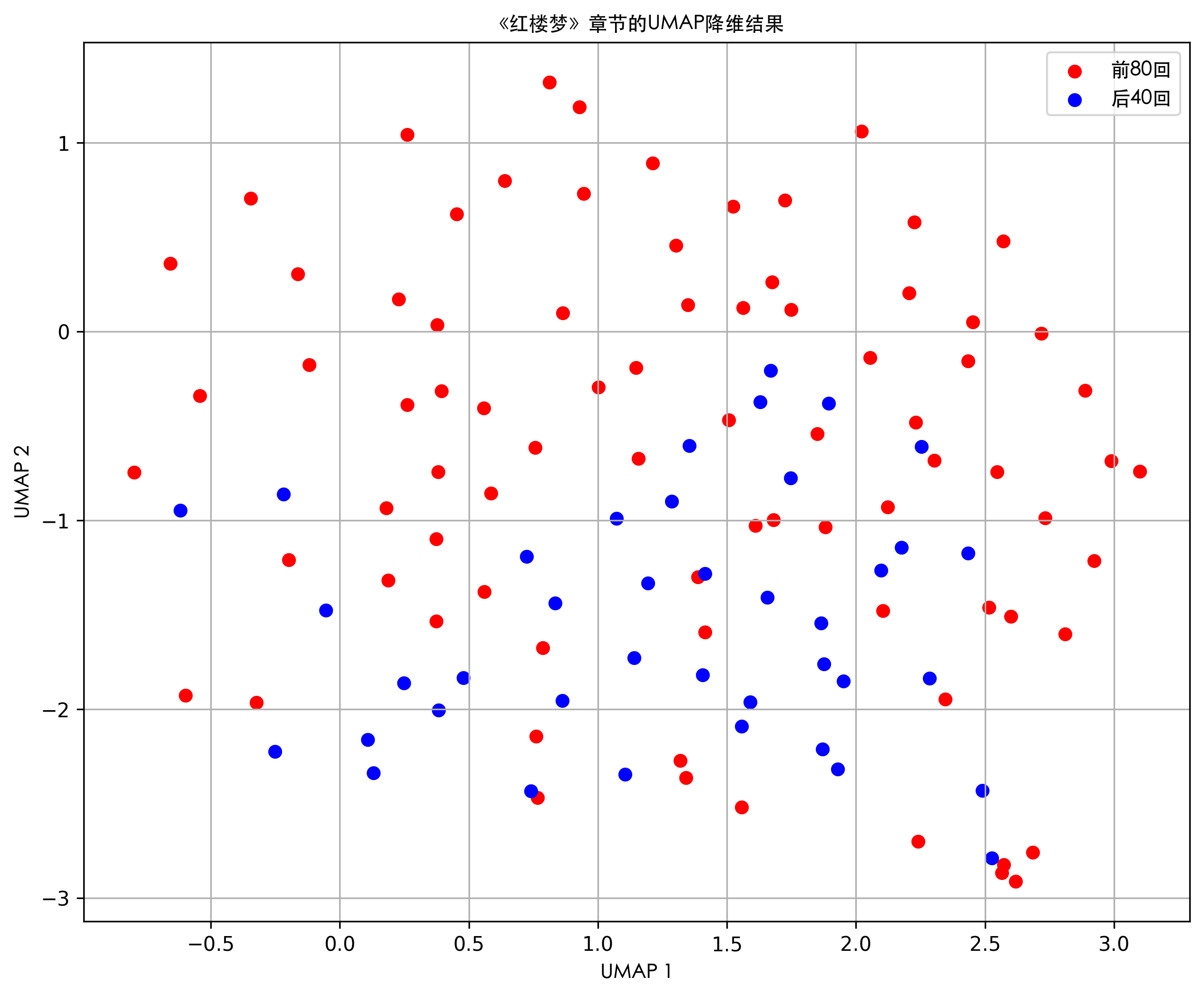
3. pca



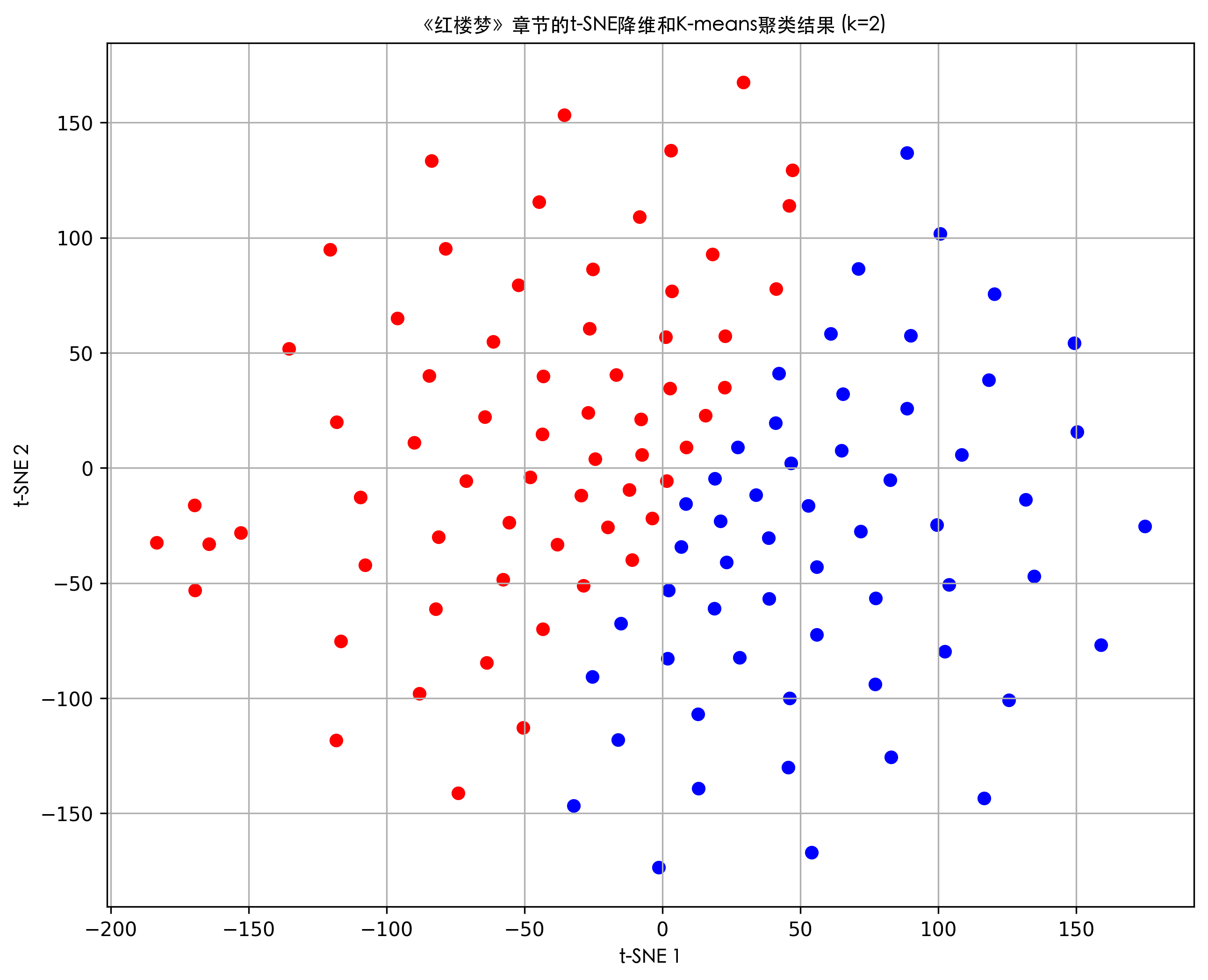
4. tsne



5. umap



6. kmeans

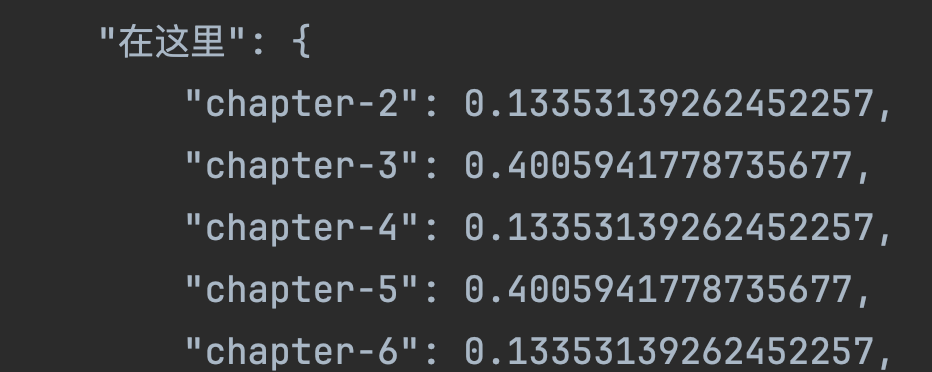


c. 3-gram

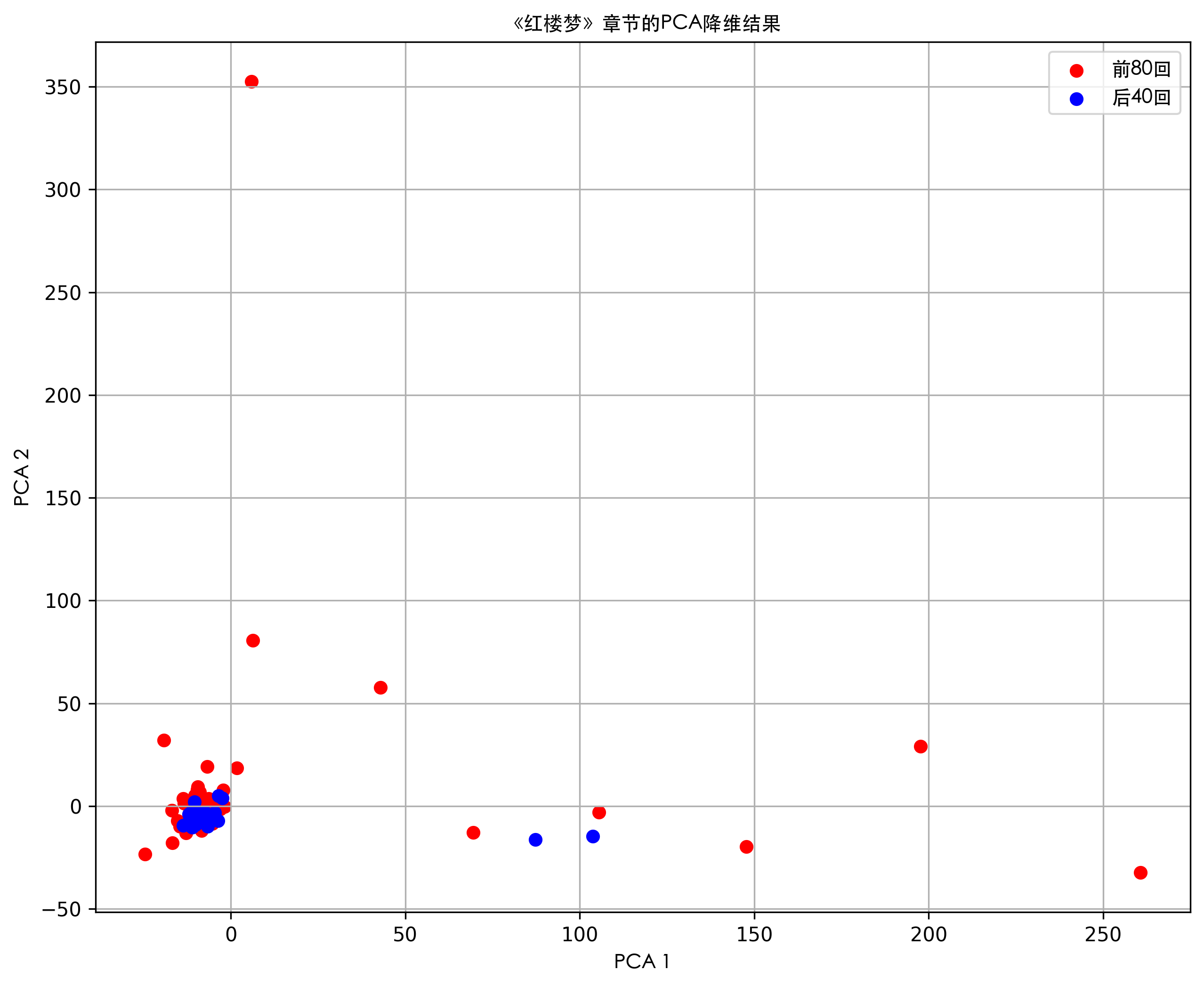
1. 数据预处理



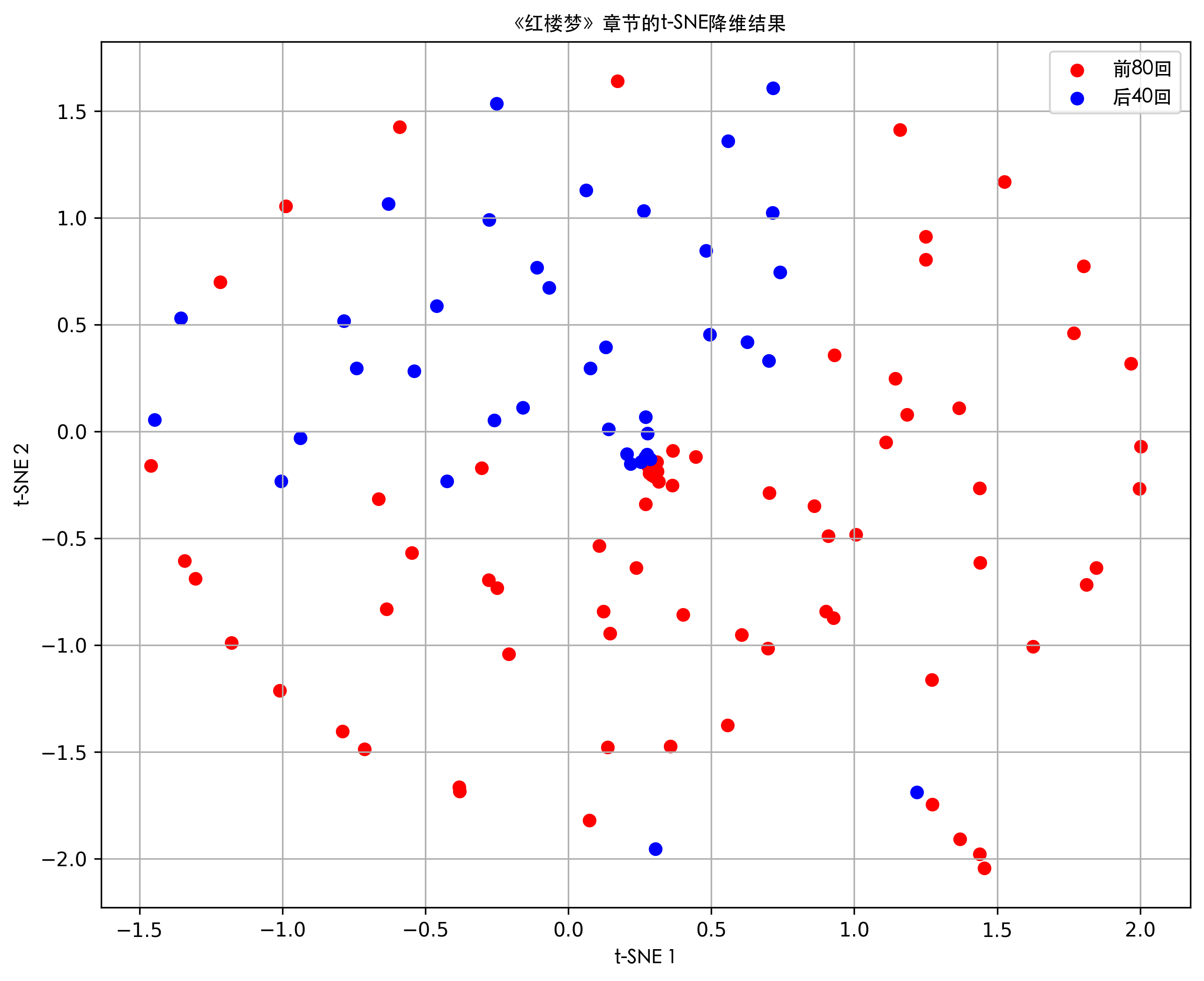
2. tfidf



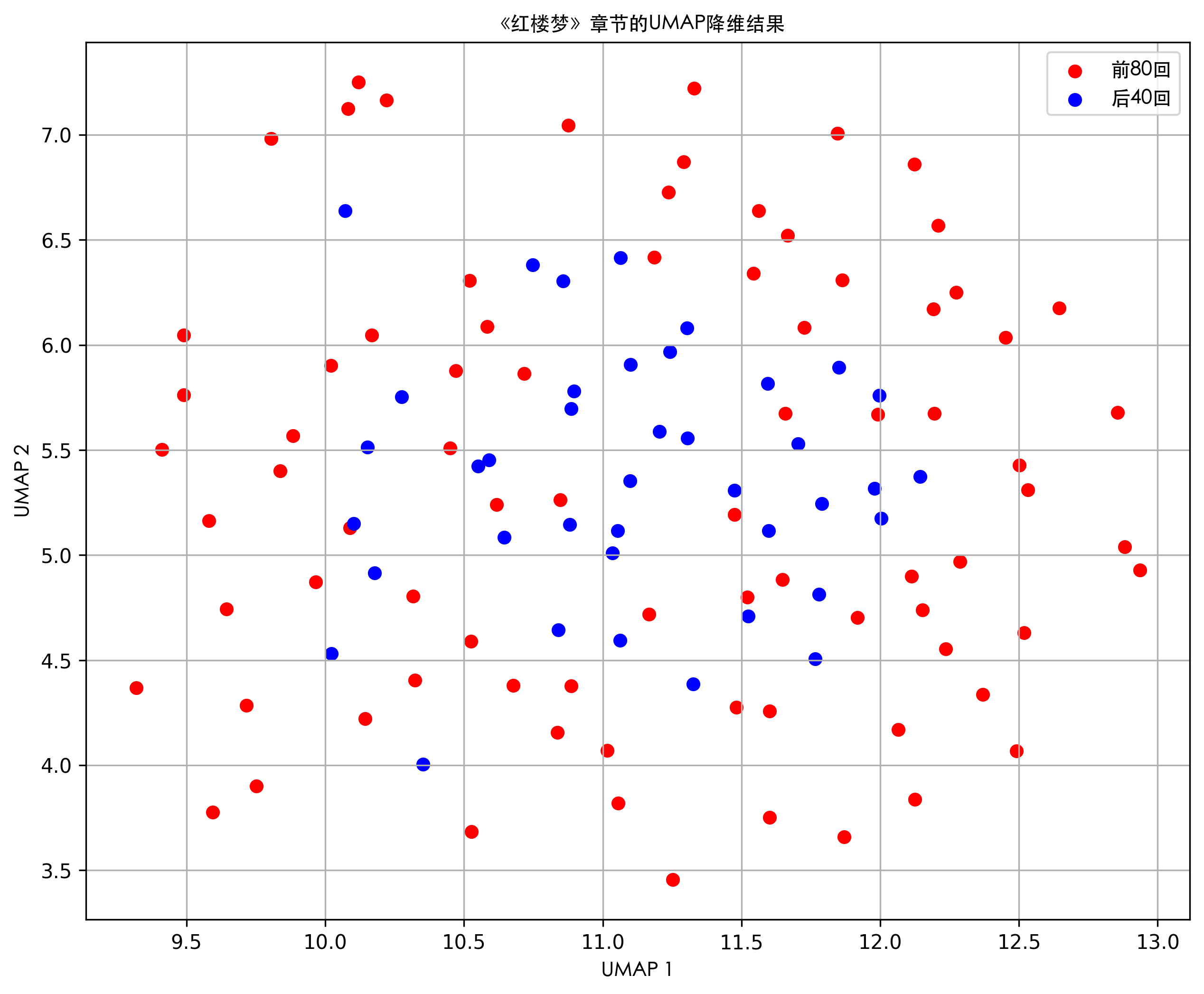
3. pca



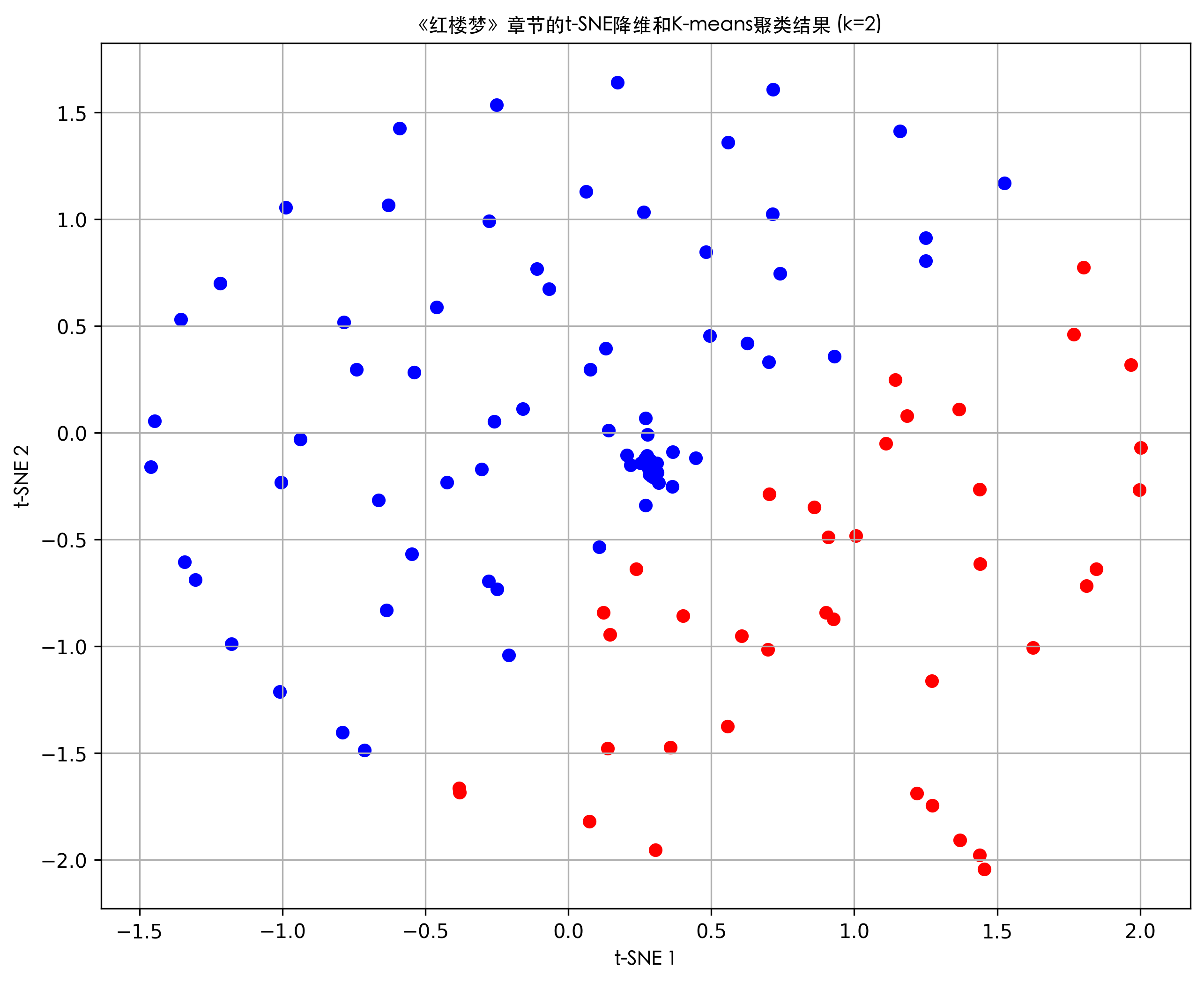
4. tsne



5. umap



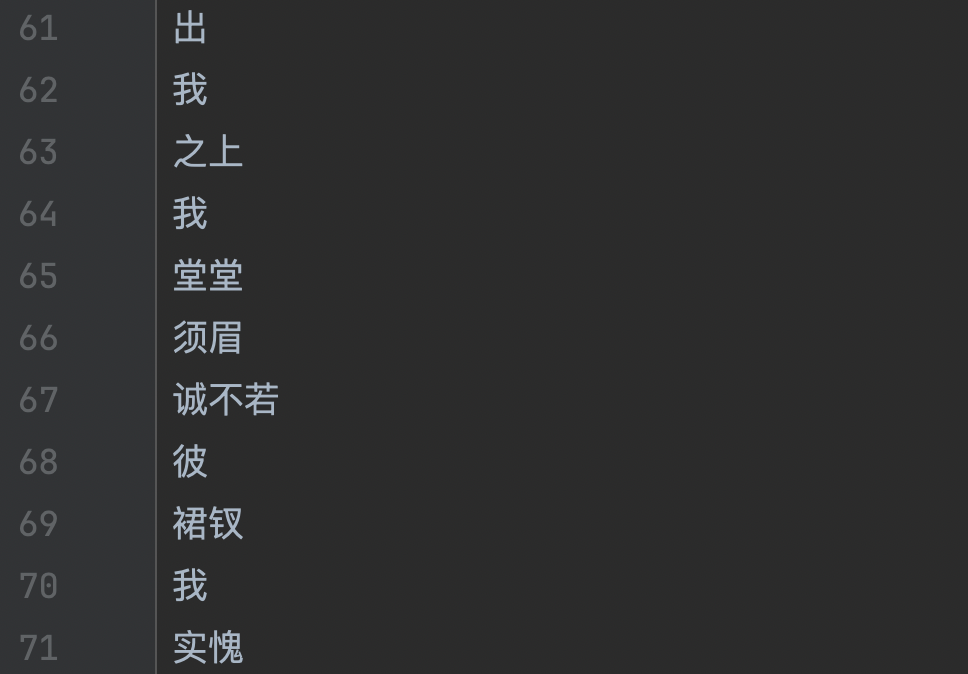
6. kmeans



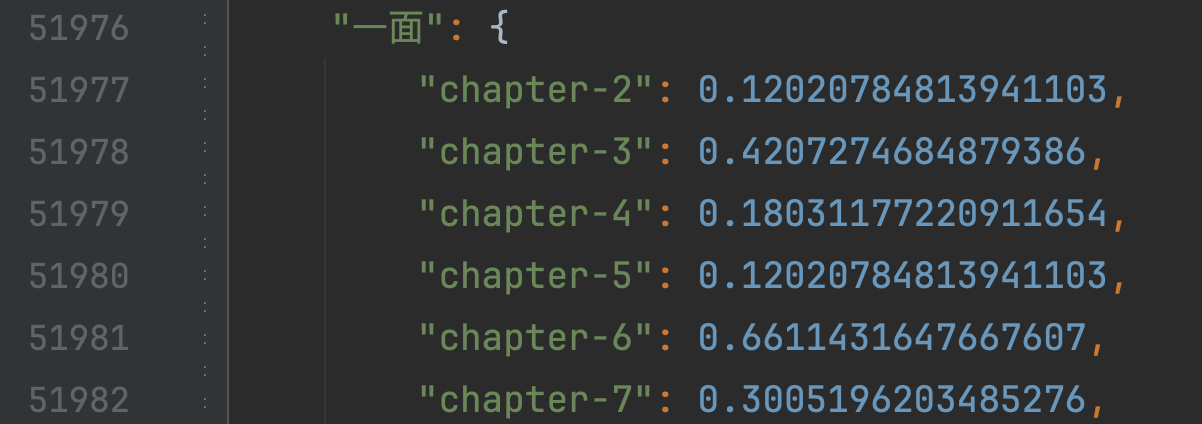
二、 按词

a. 1-gram

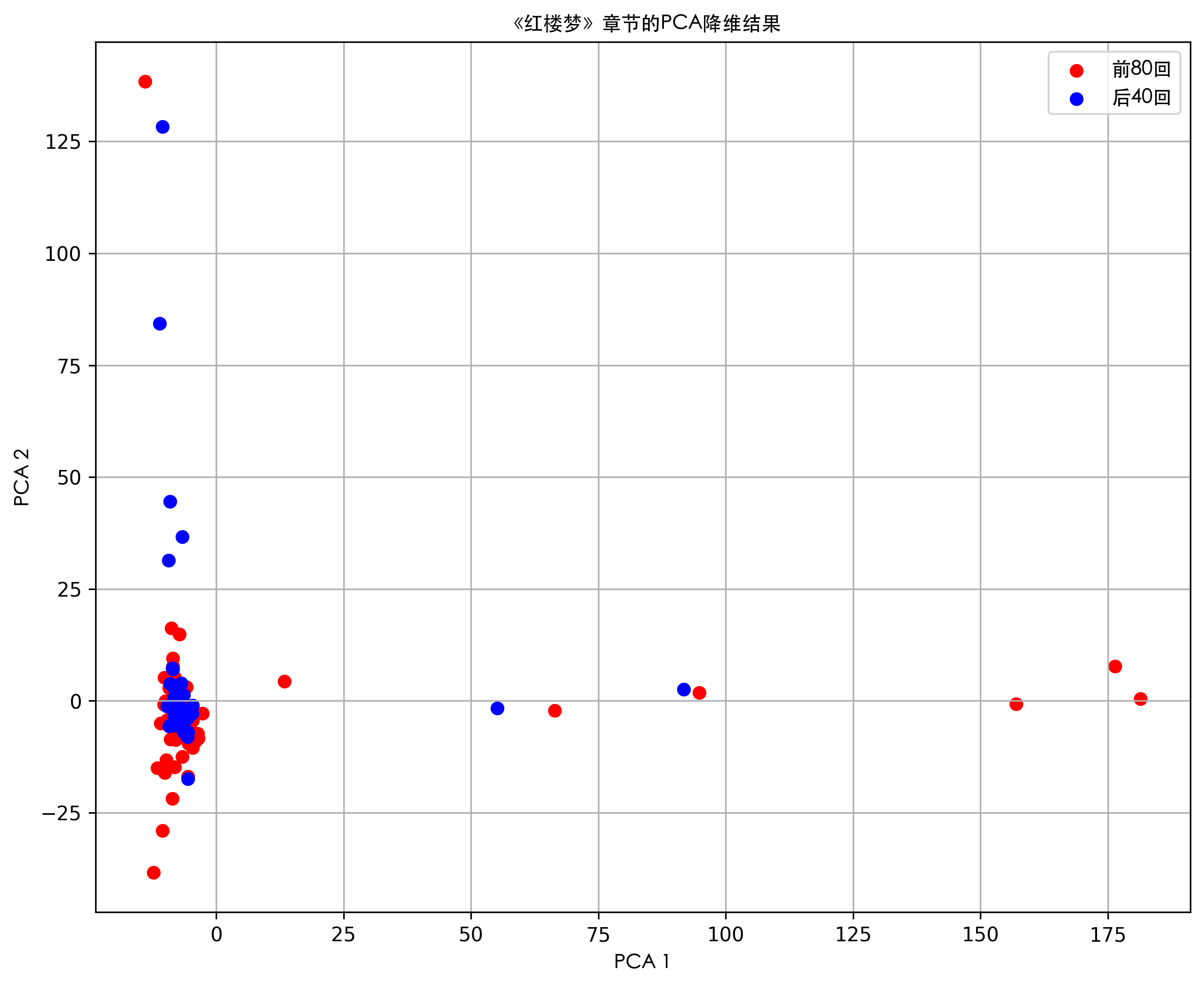
1. 数据预处理



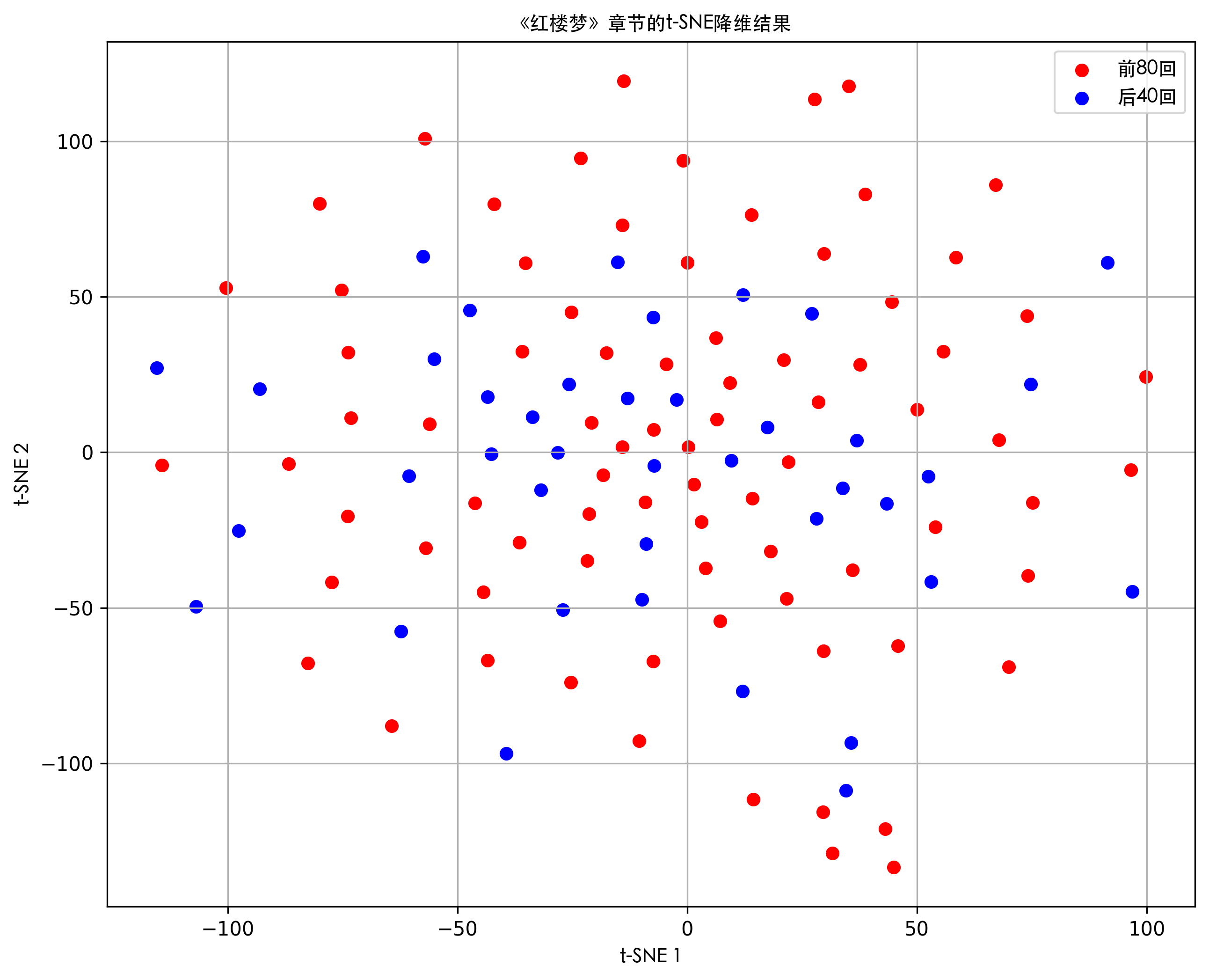
2. tfidf



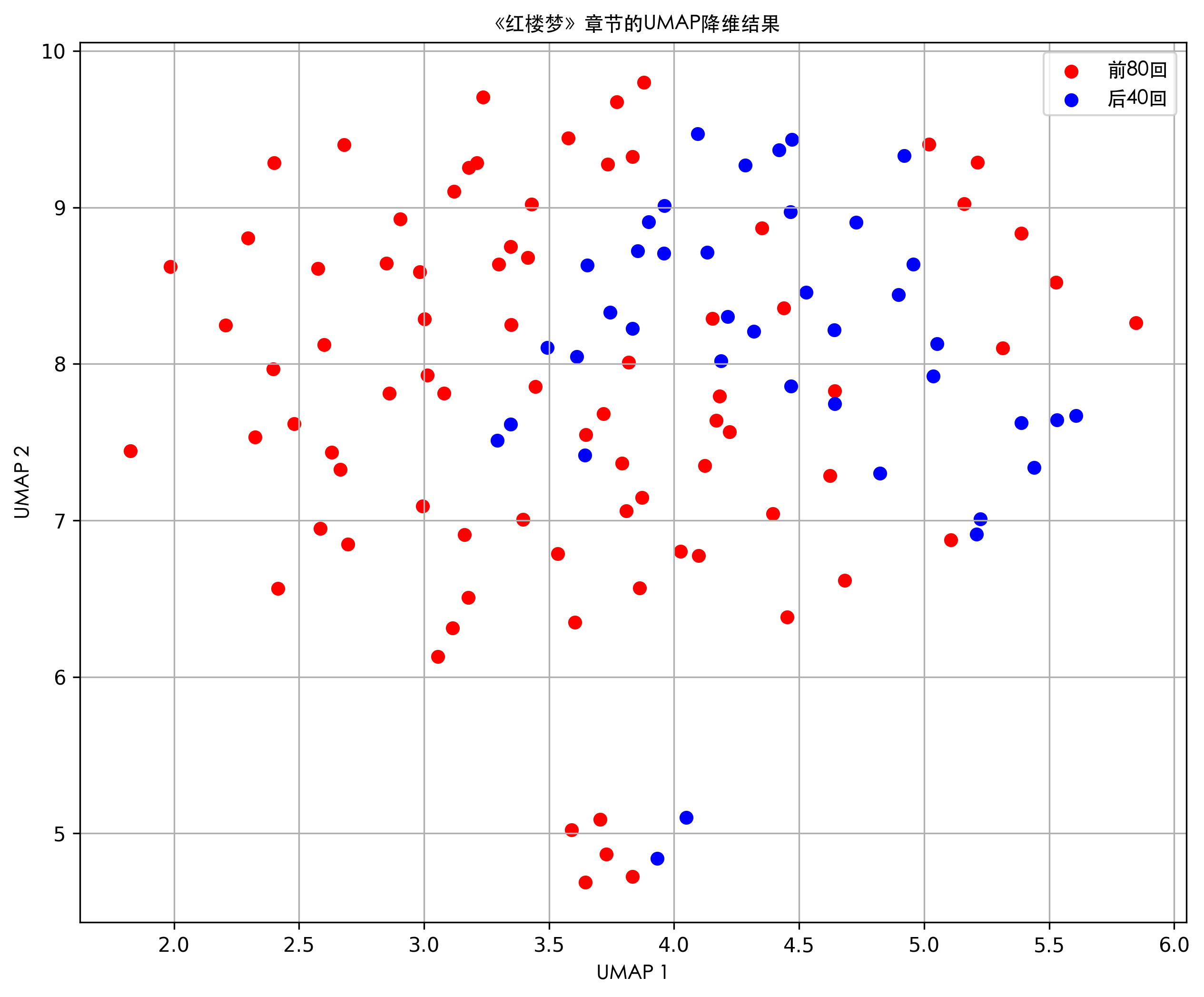
3. pca



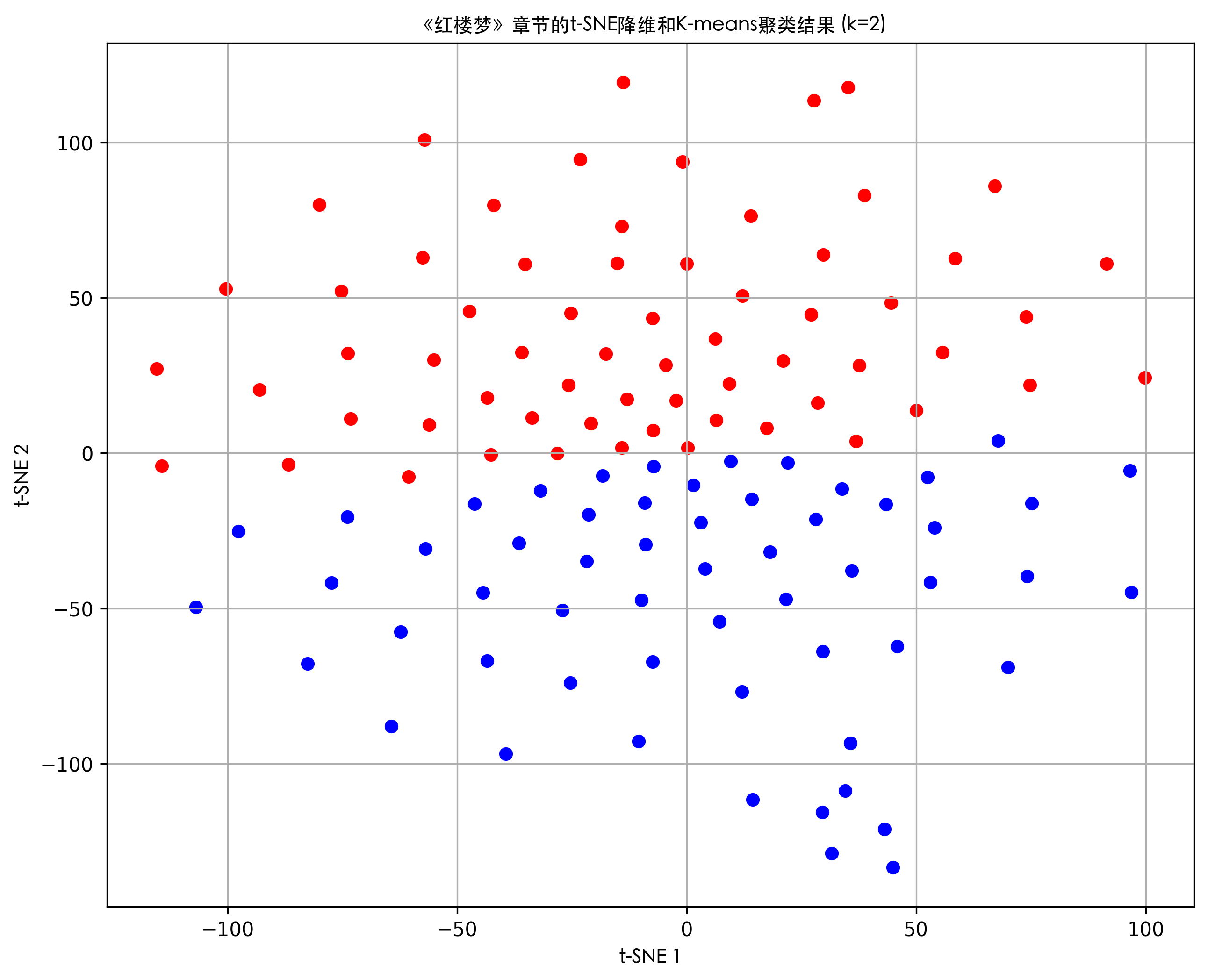
4. tsne



5. umap

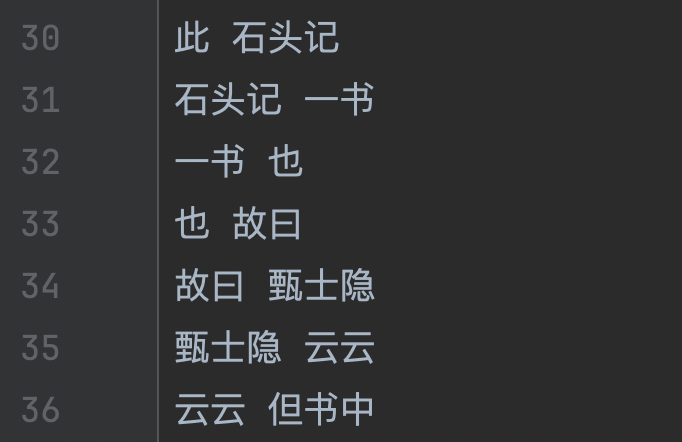


6. kmeans

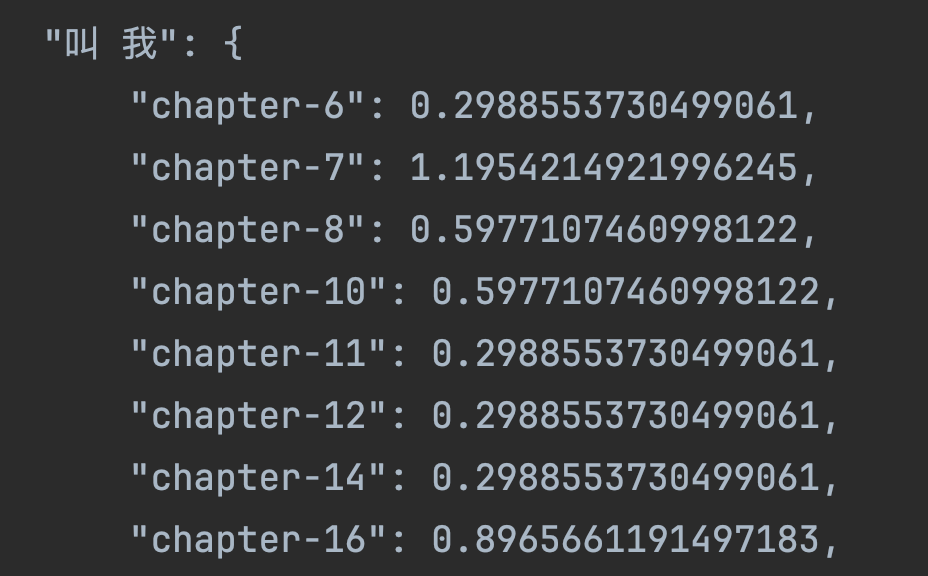


b. 2-gram

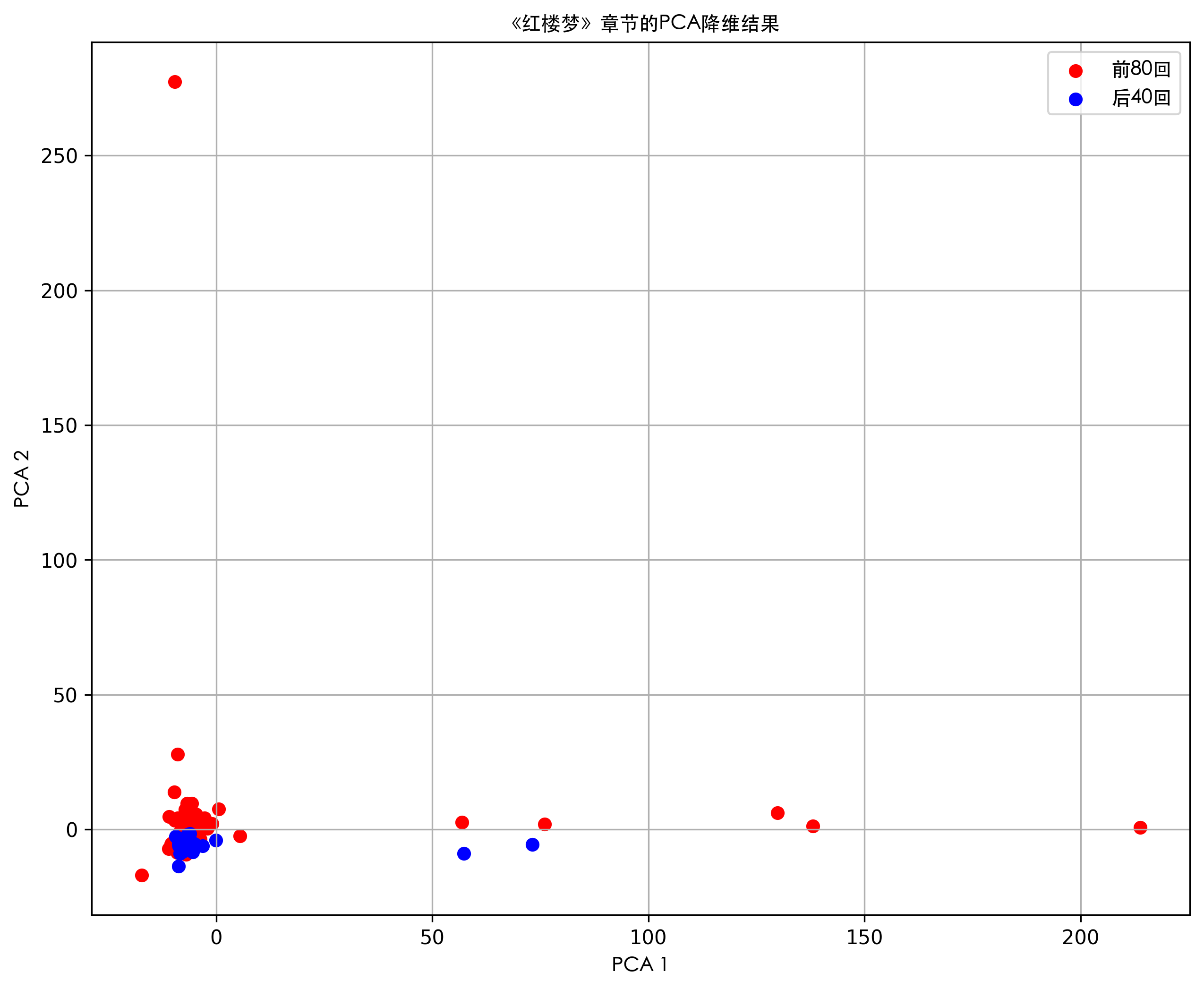
1. 数据预处理



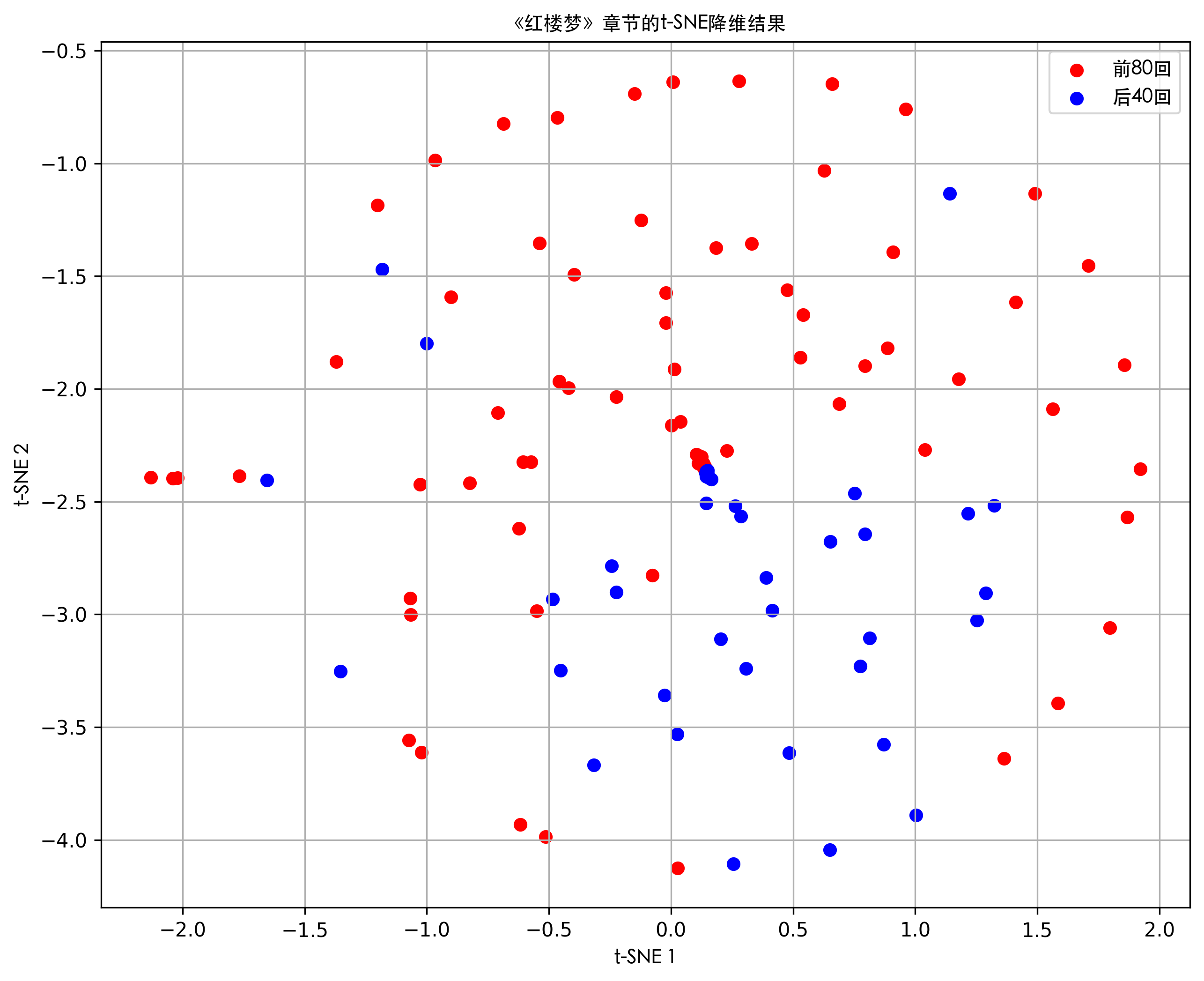
2. tfidf



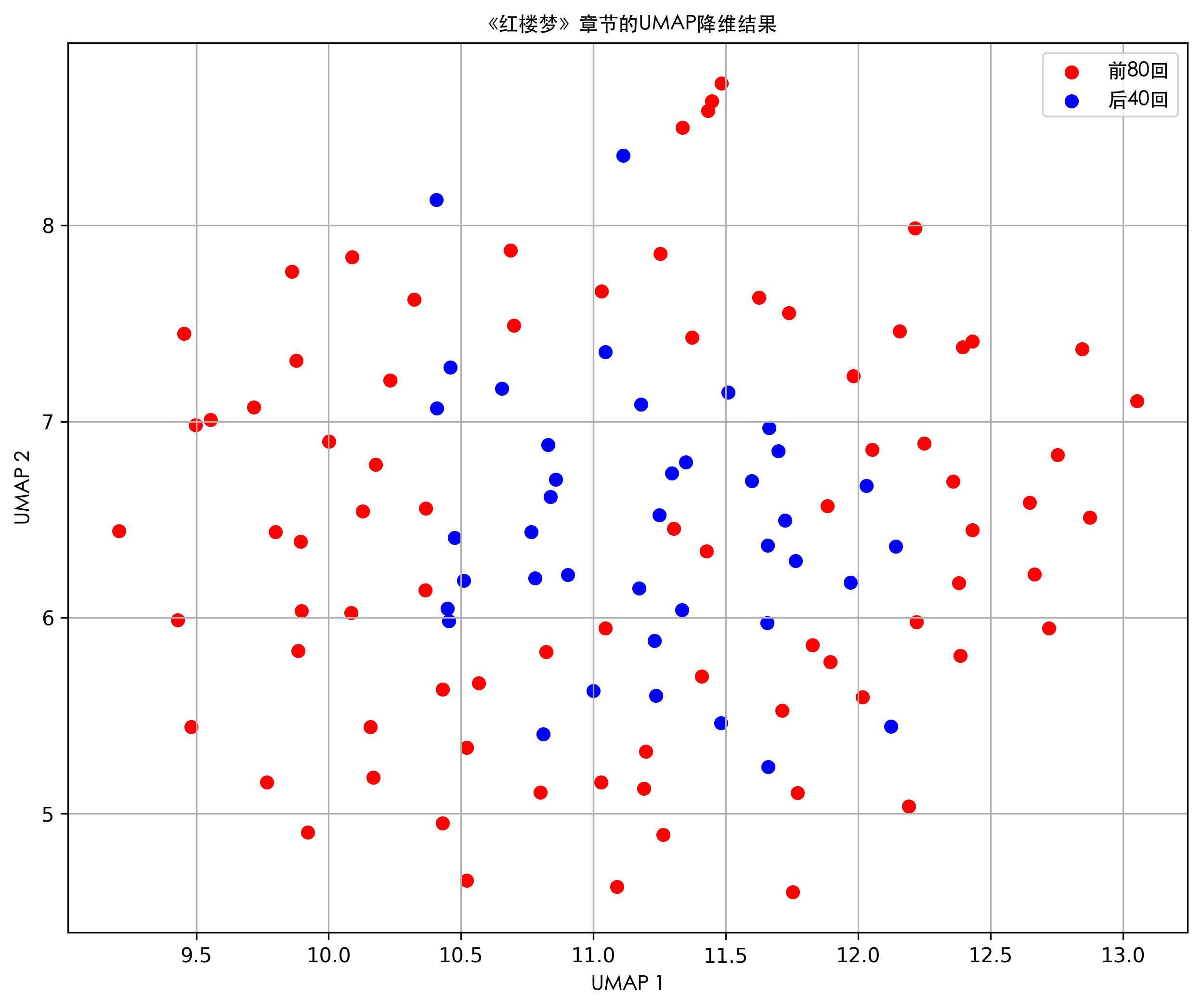
3. pca



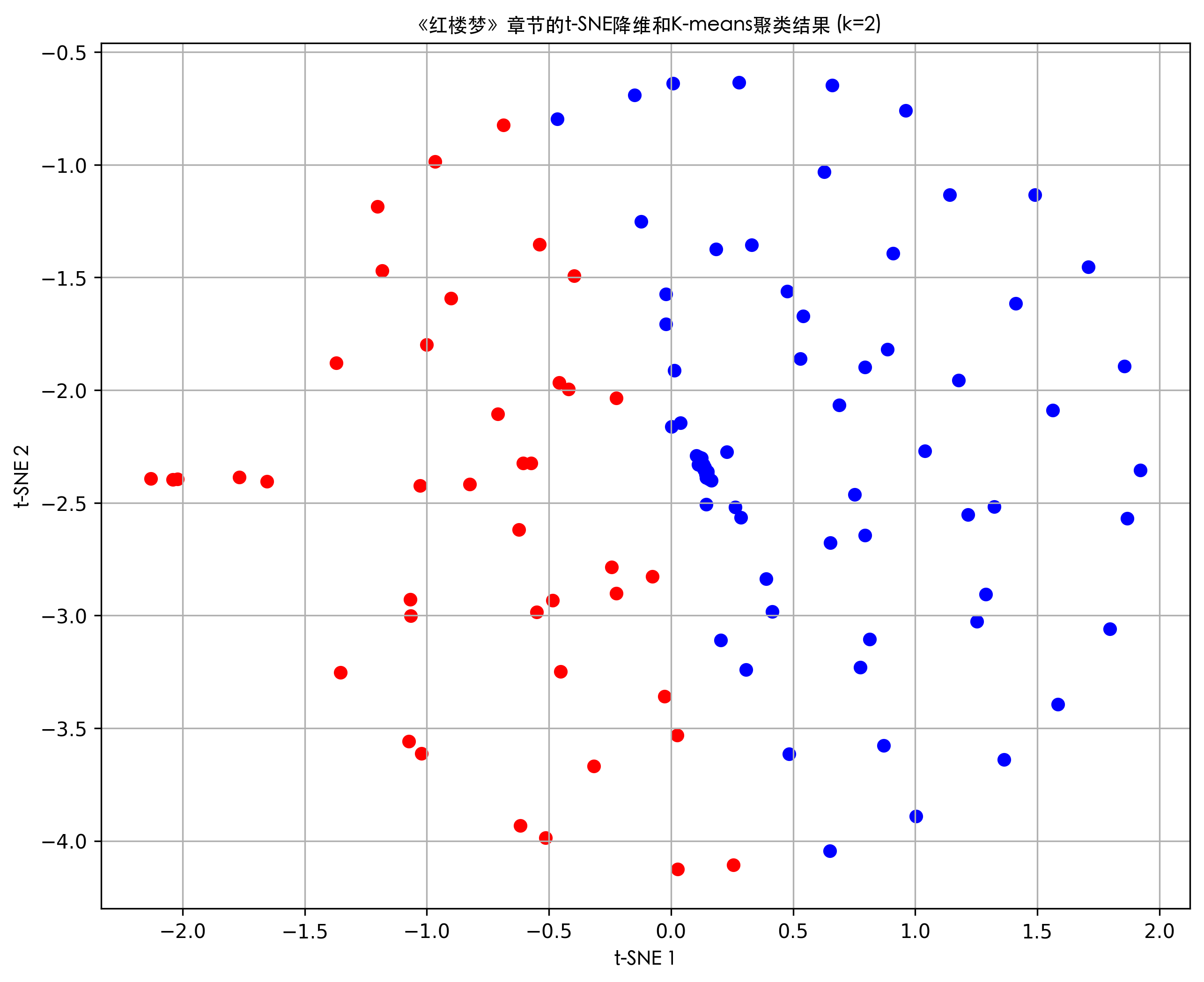
4. tsne



5. umap

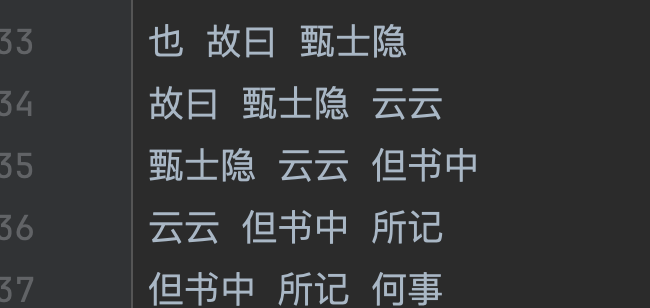


6. kmeans

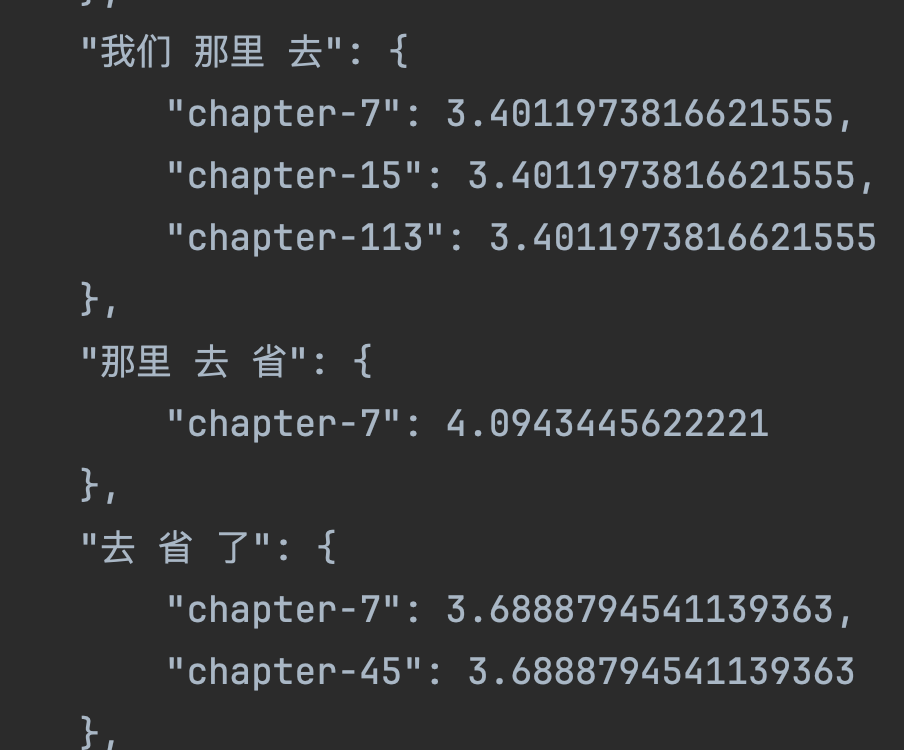


c. 3-gram

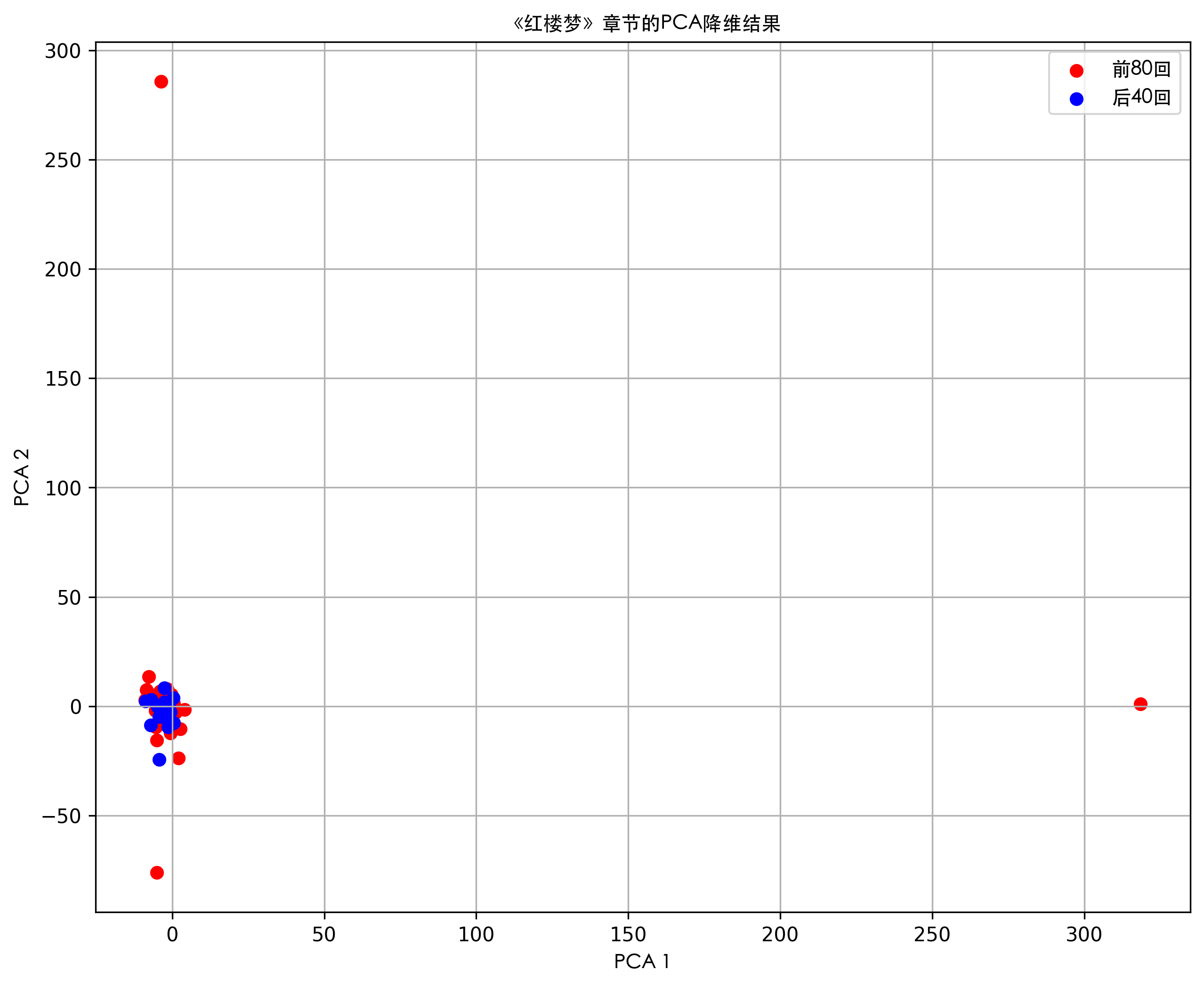
1. 数据预处理



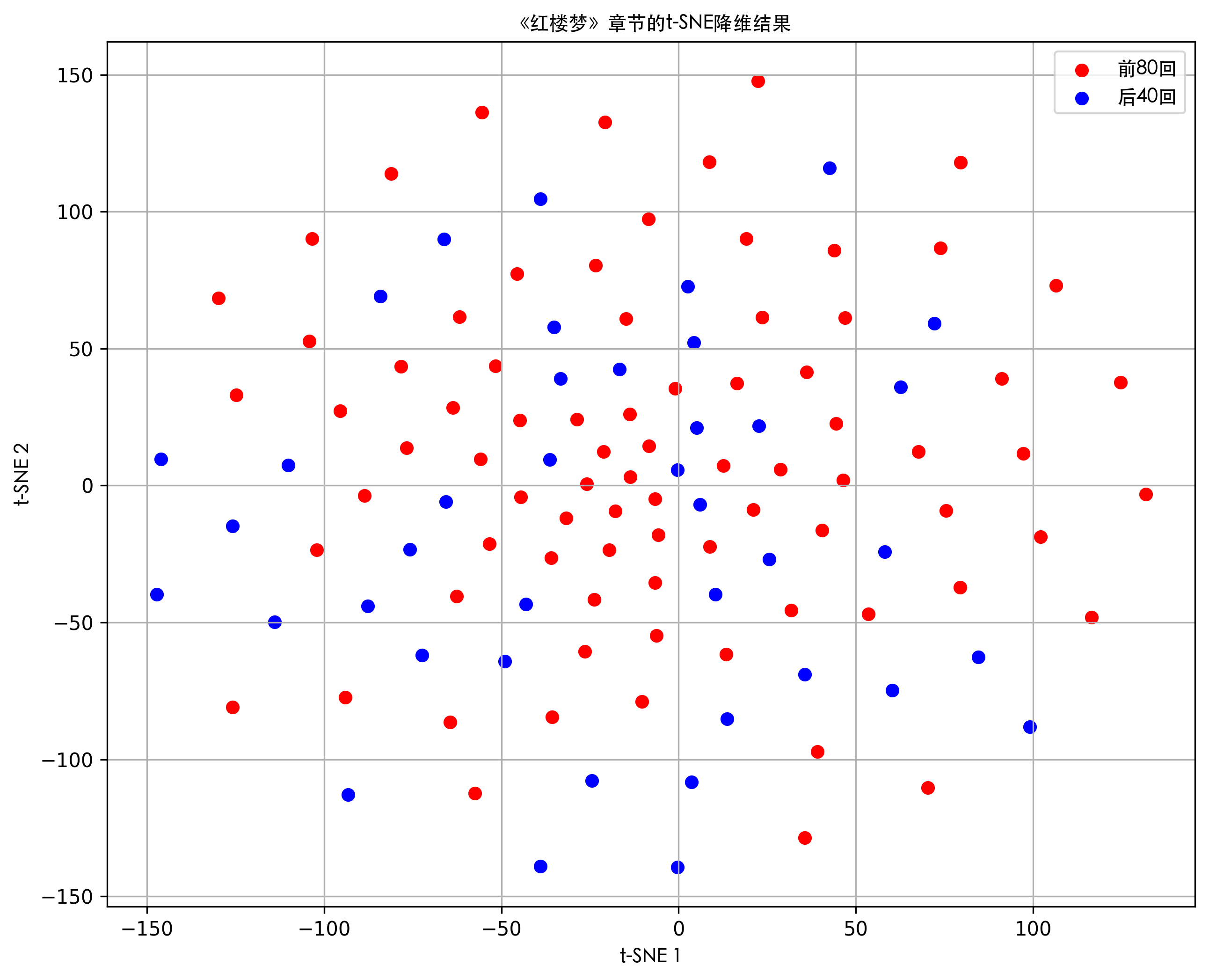
2. tfidf



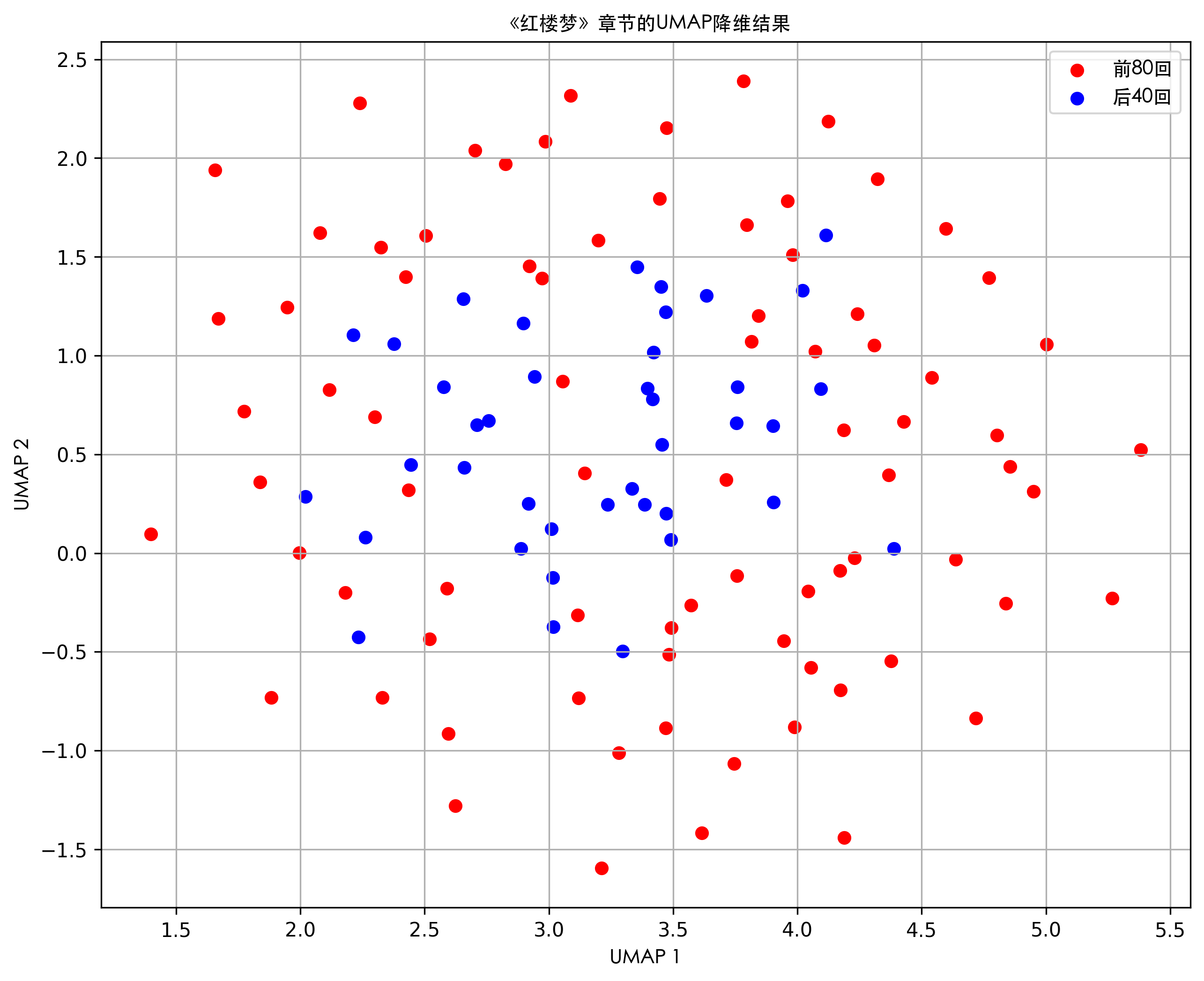
3. pca



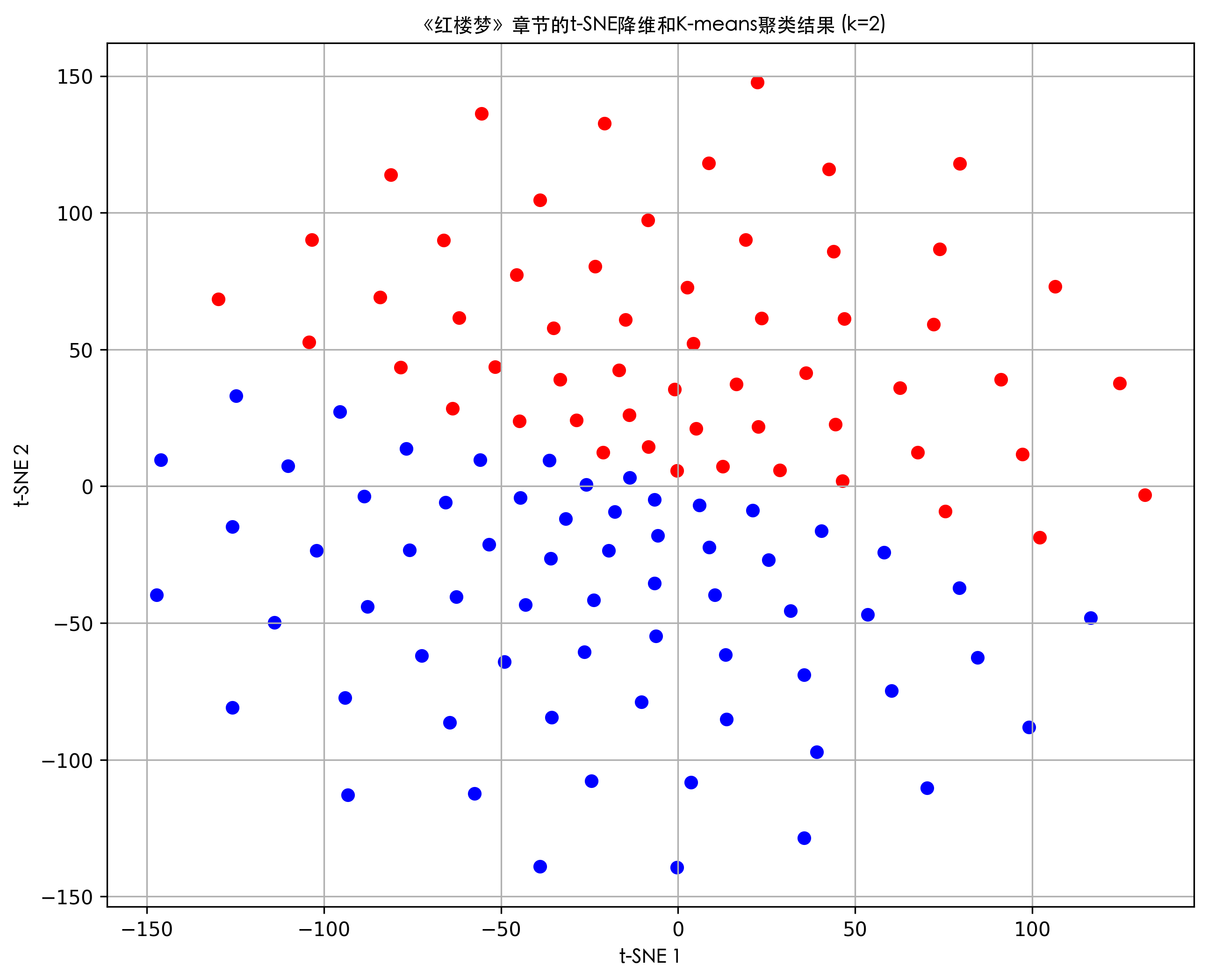
4. tsne



5. umap



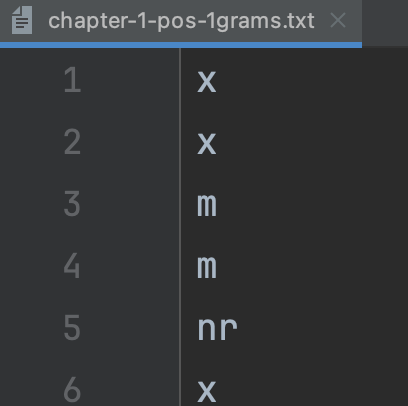
6. kmeans



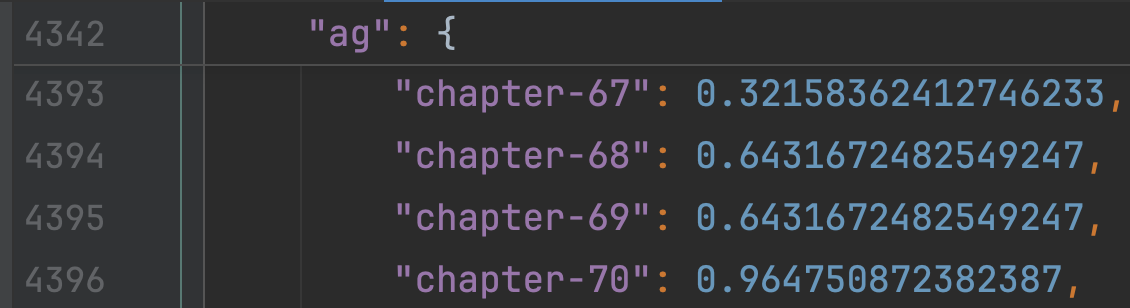
三、 按词性

a. 1-gram

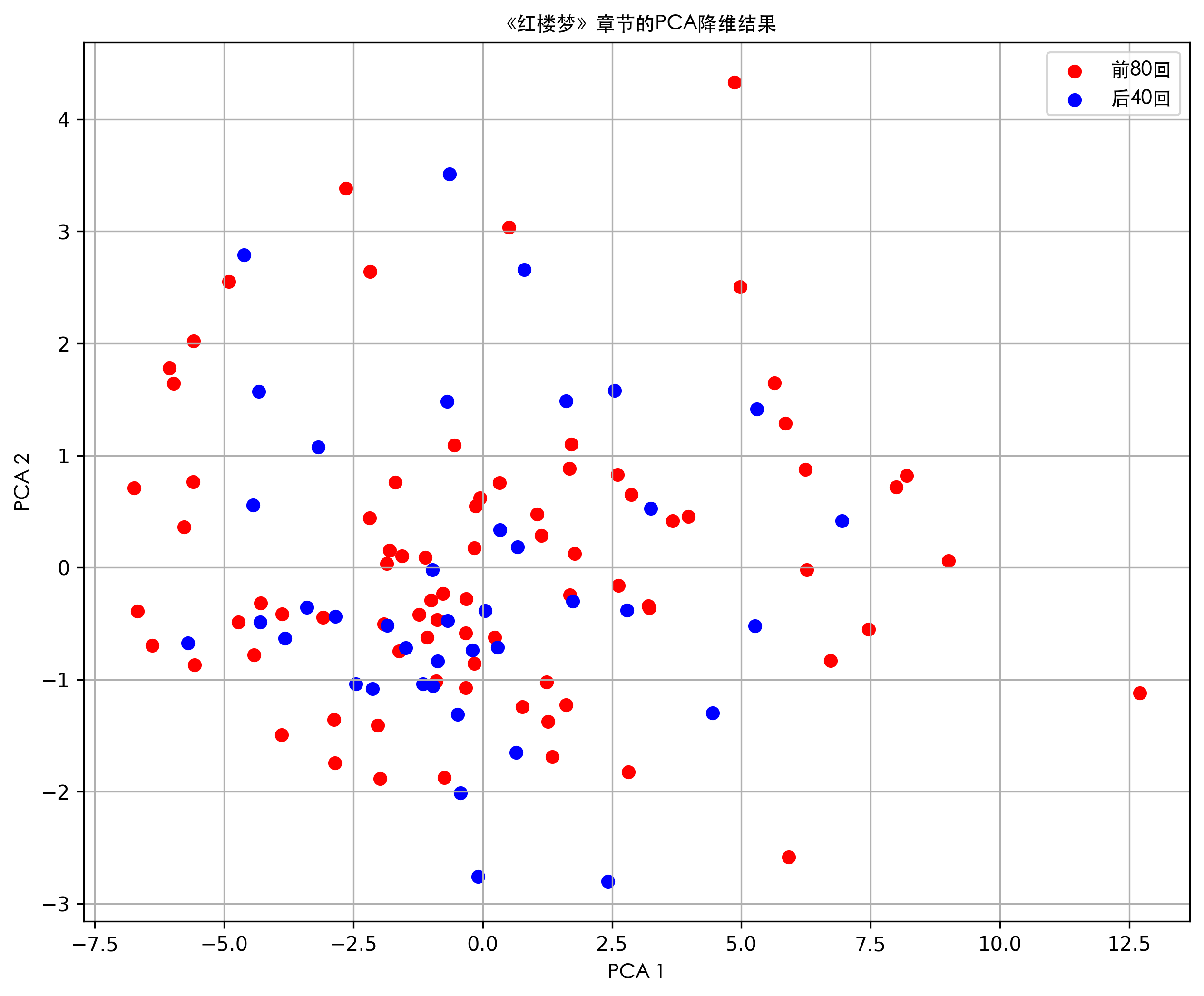
1. 数据预处理



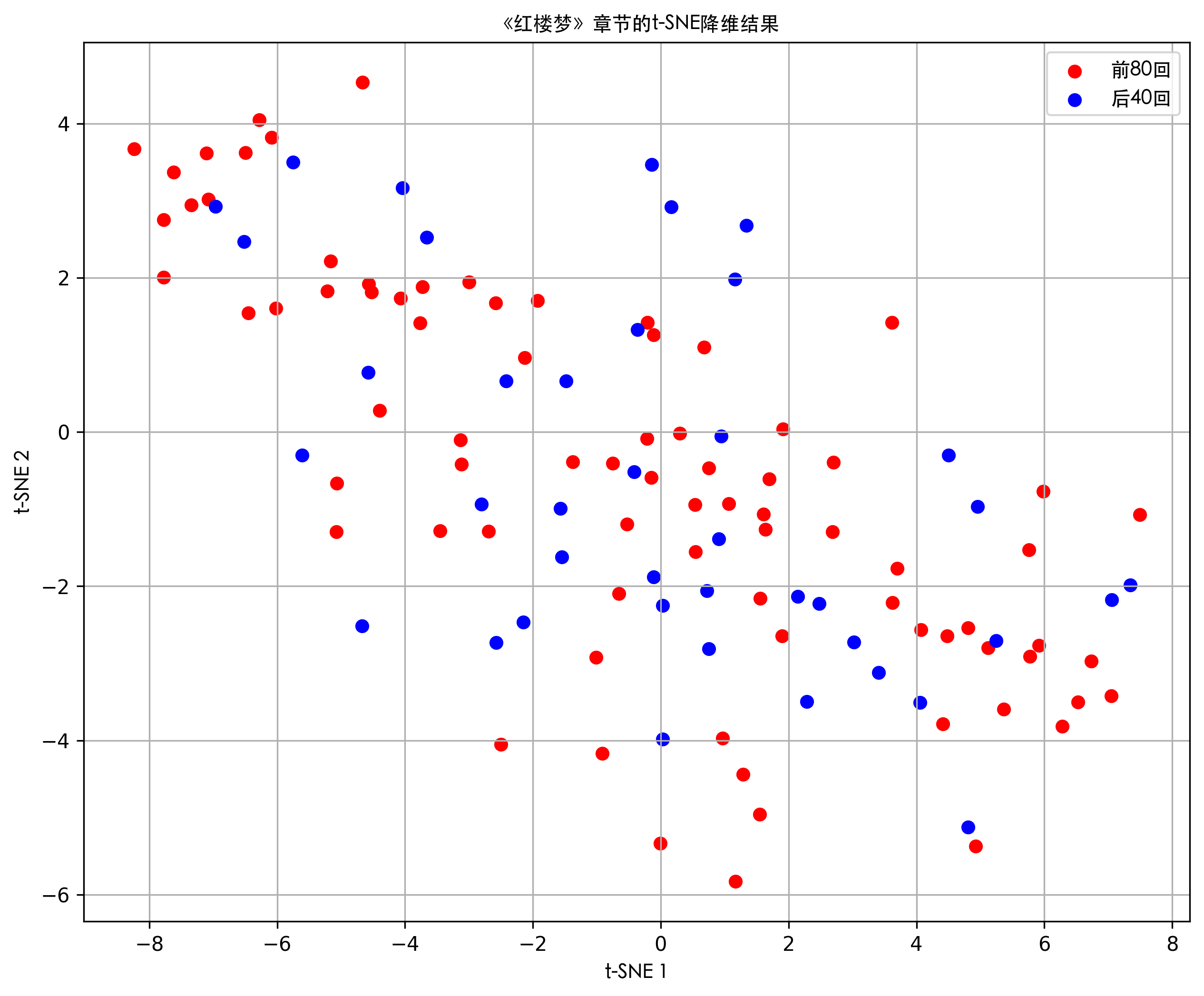
2. tfidf



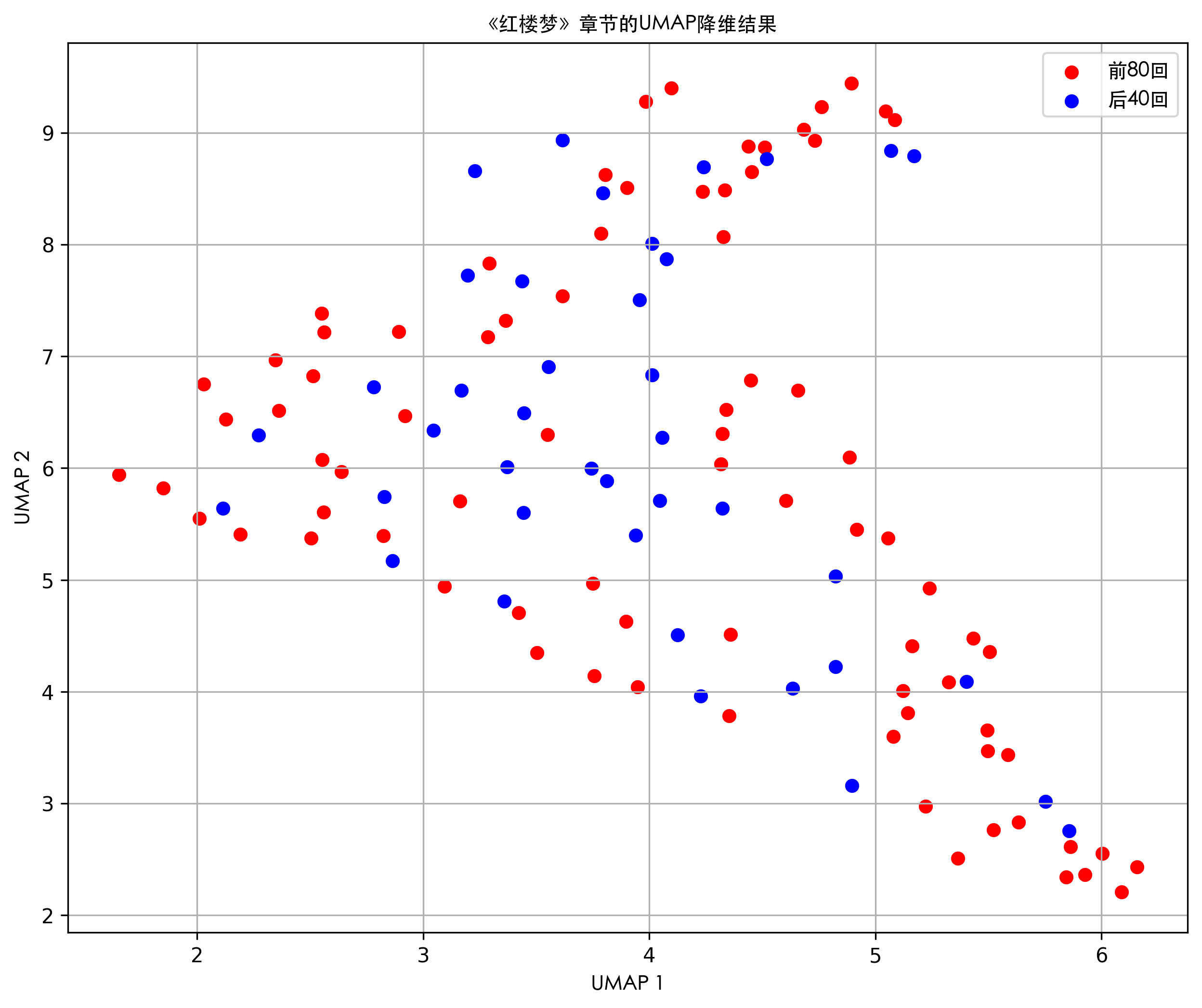
3. pca



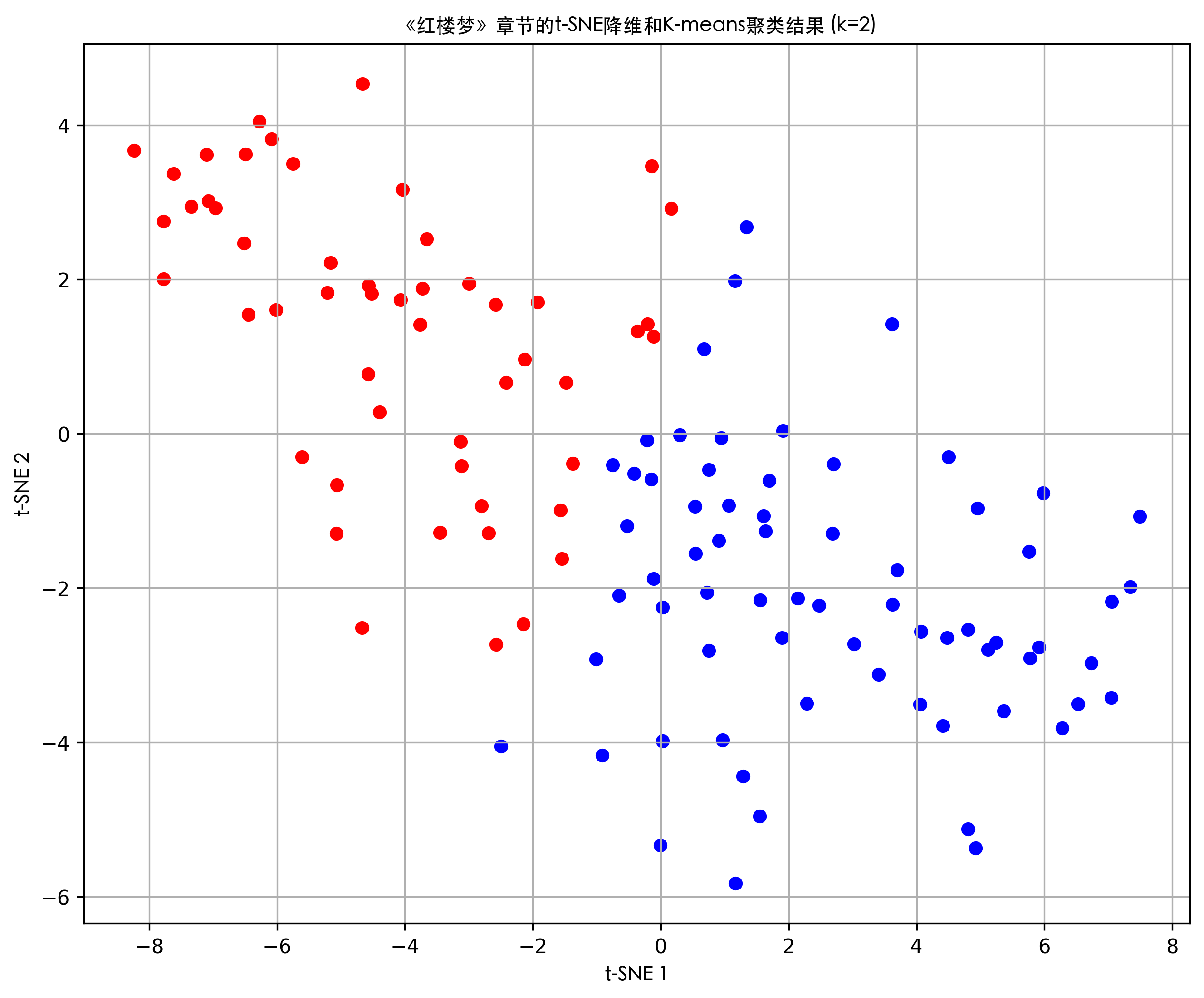
4. tsne



5. umap

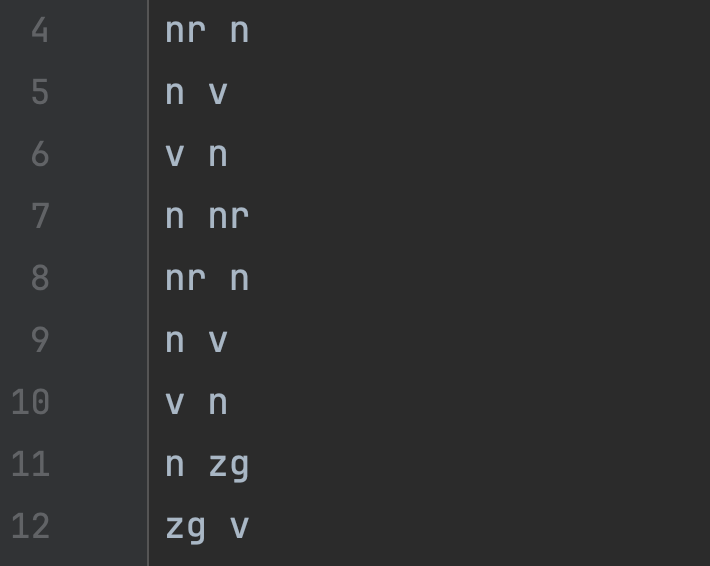


6. kmeans

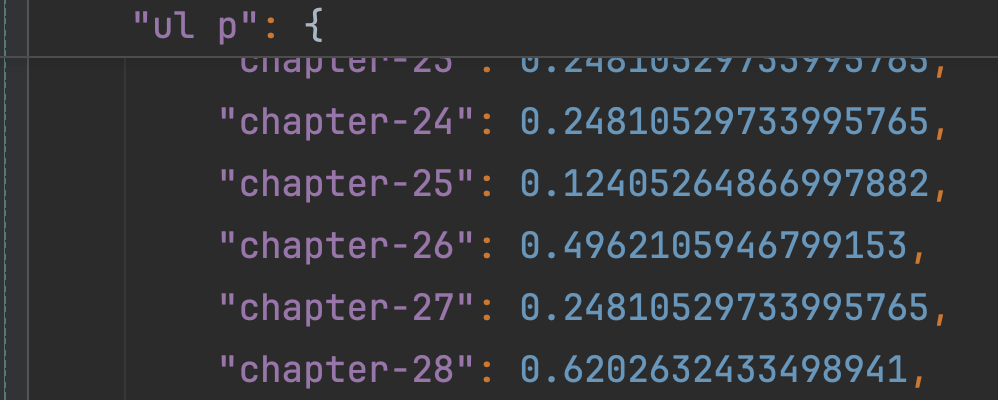


b. 2-gram

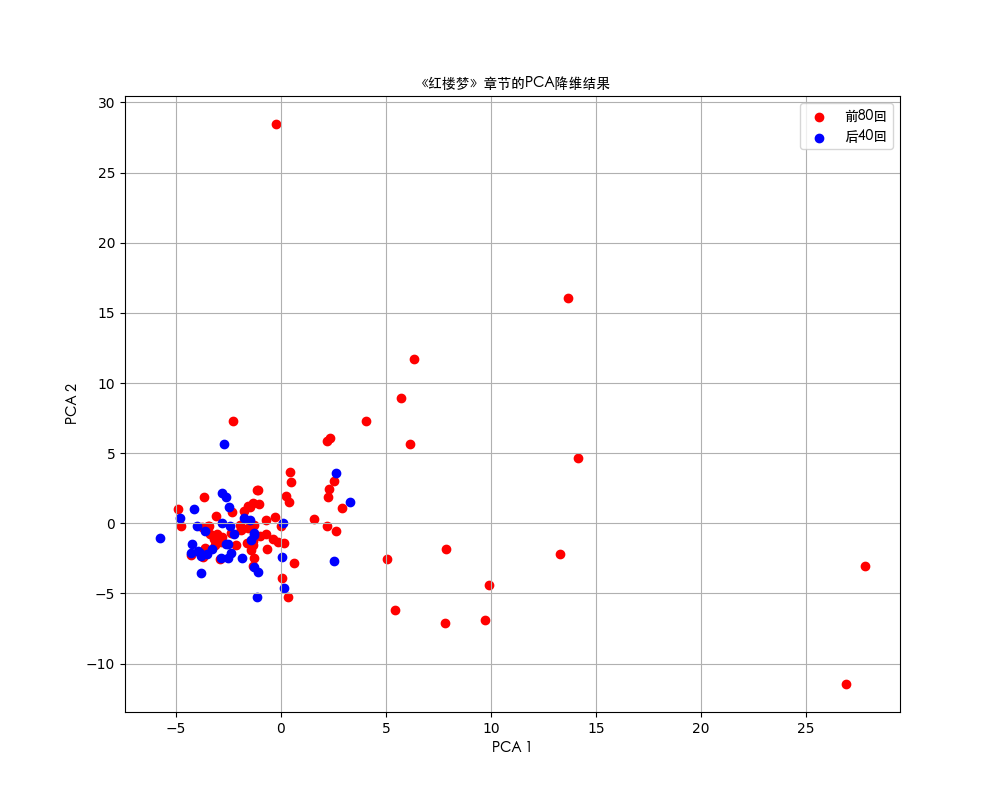
1. 数据预处理



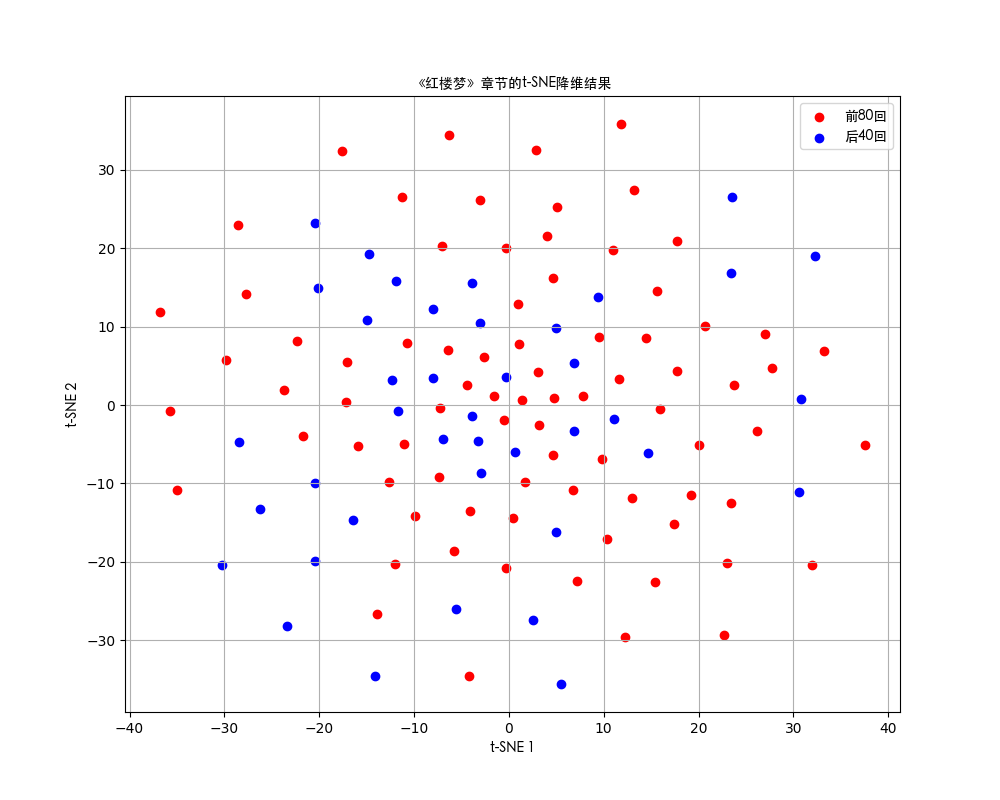
2. tfidf



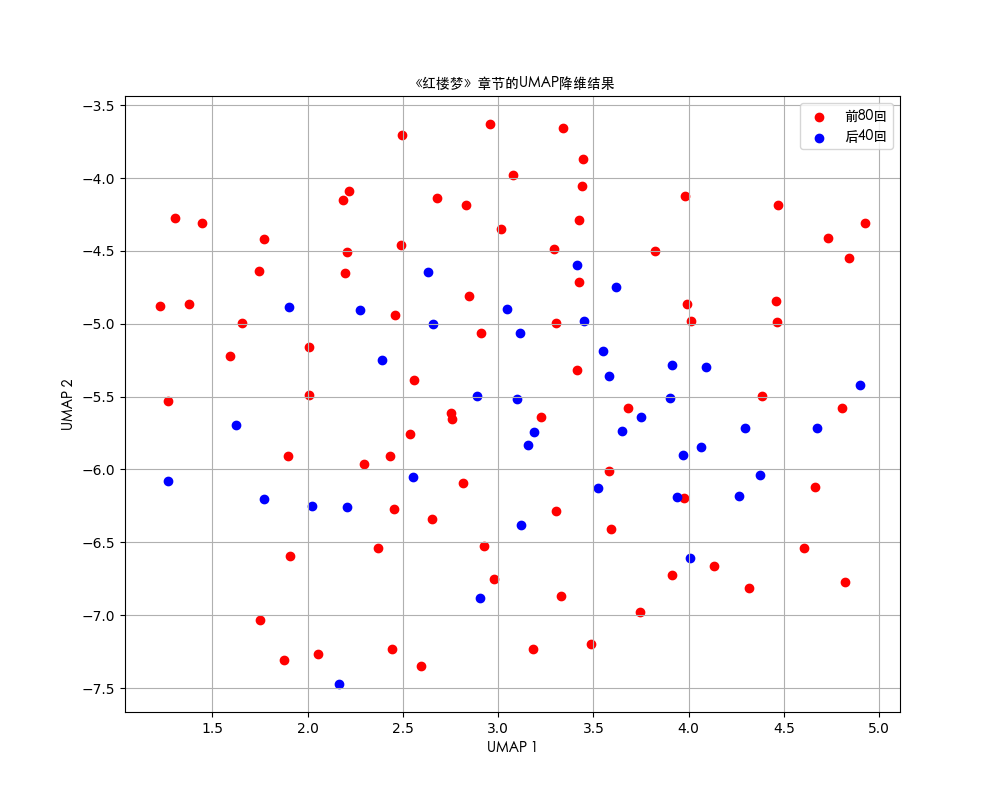
3. pca



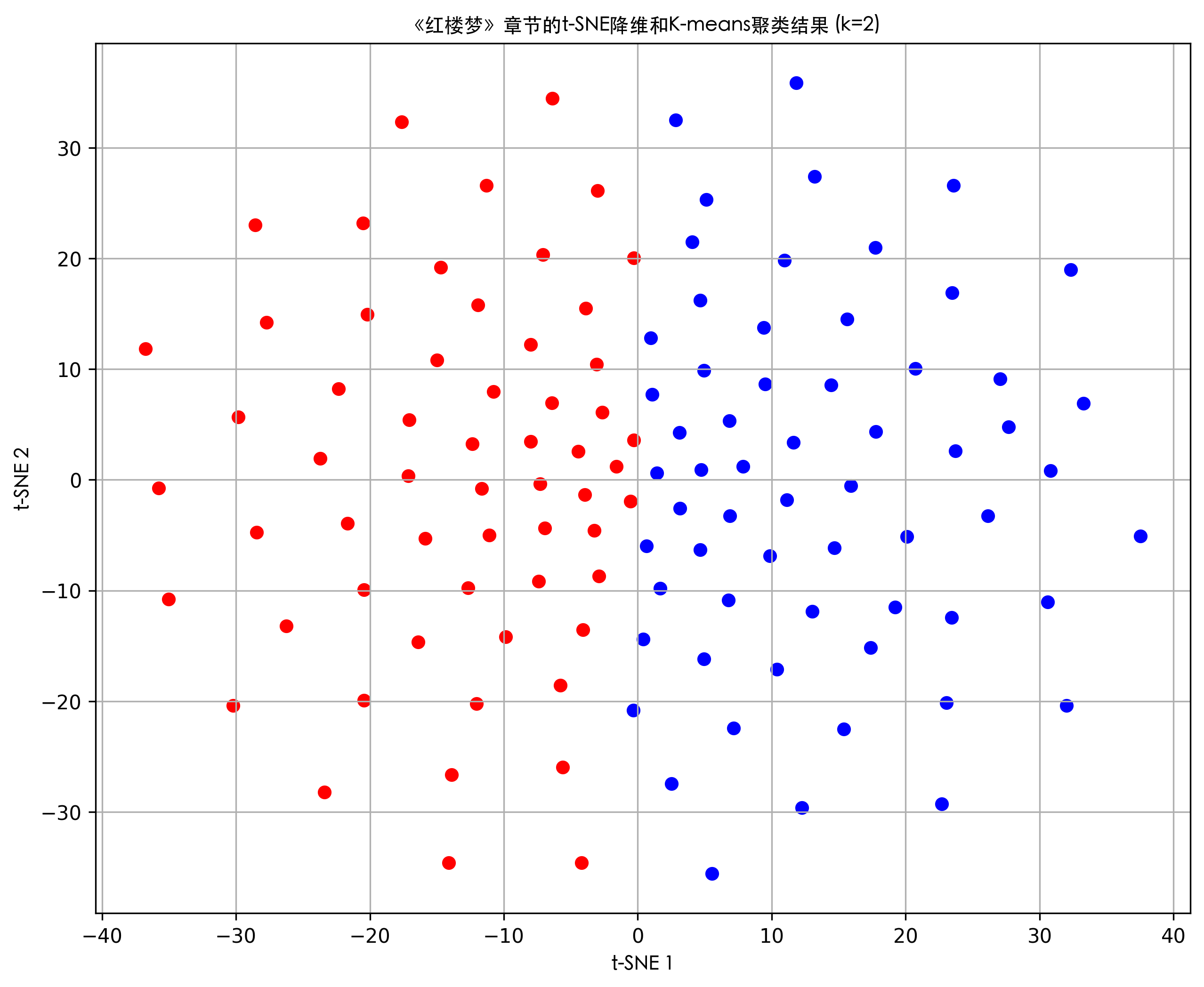
4. tsne



5. umap

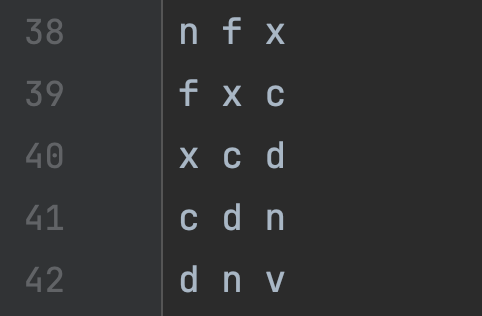


6. kmeans

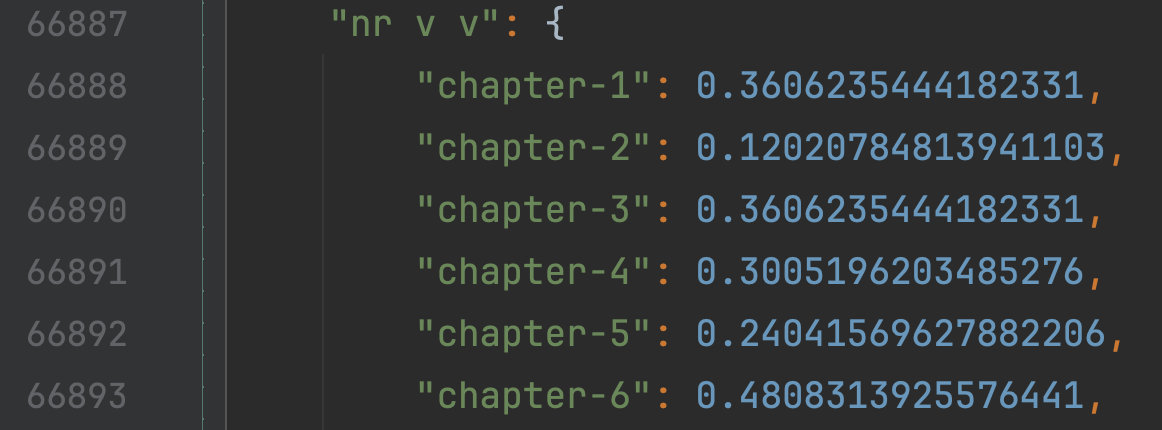


c. 3-gram

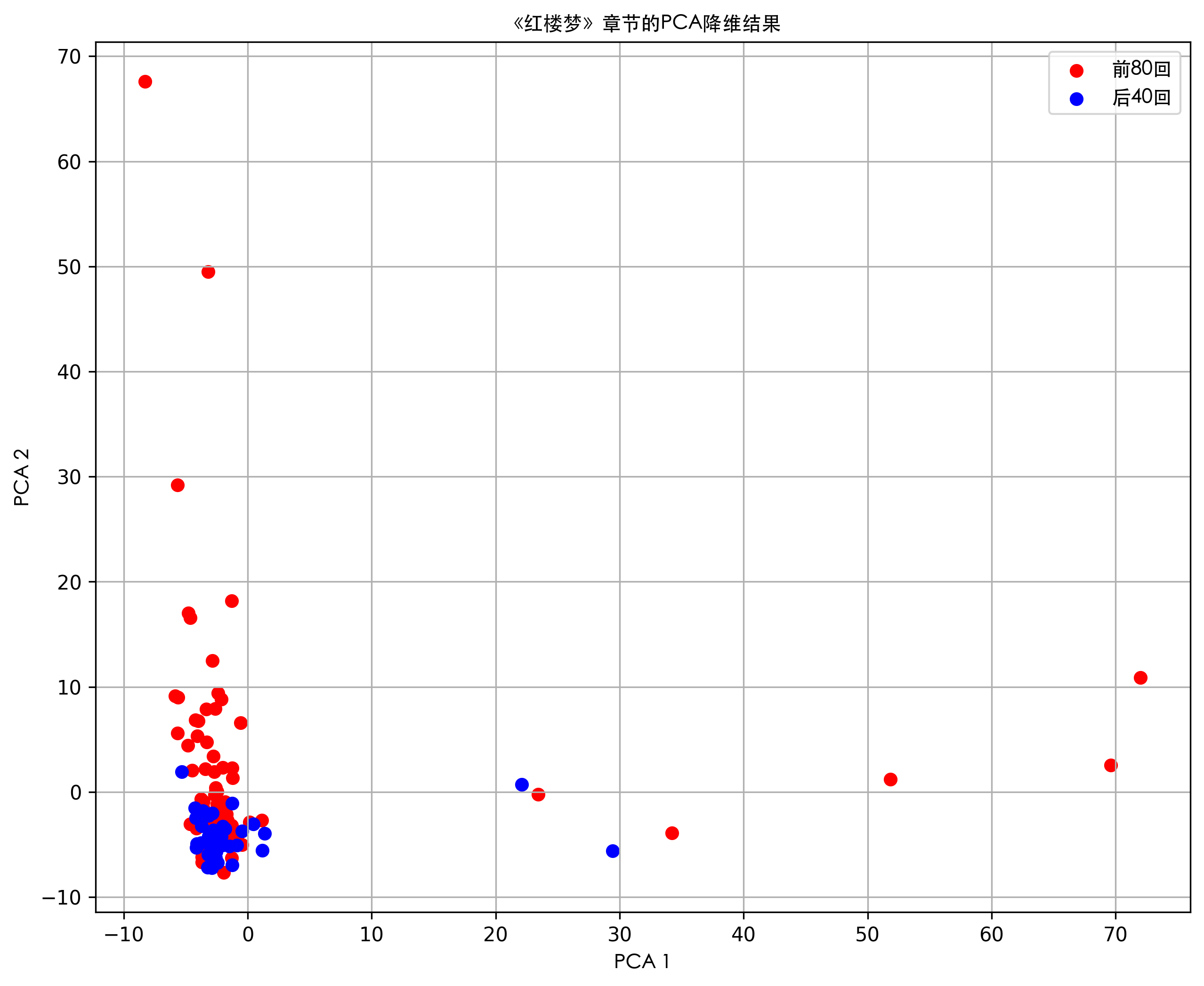
1. 数据预处理



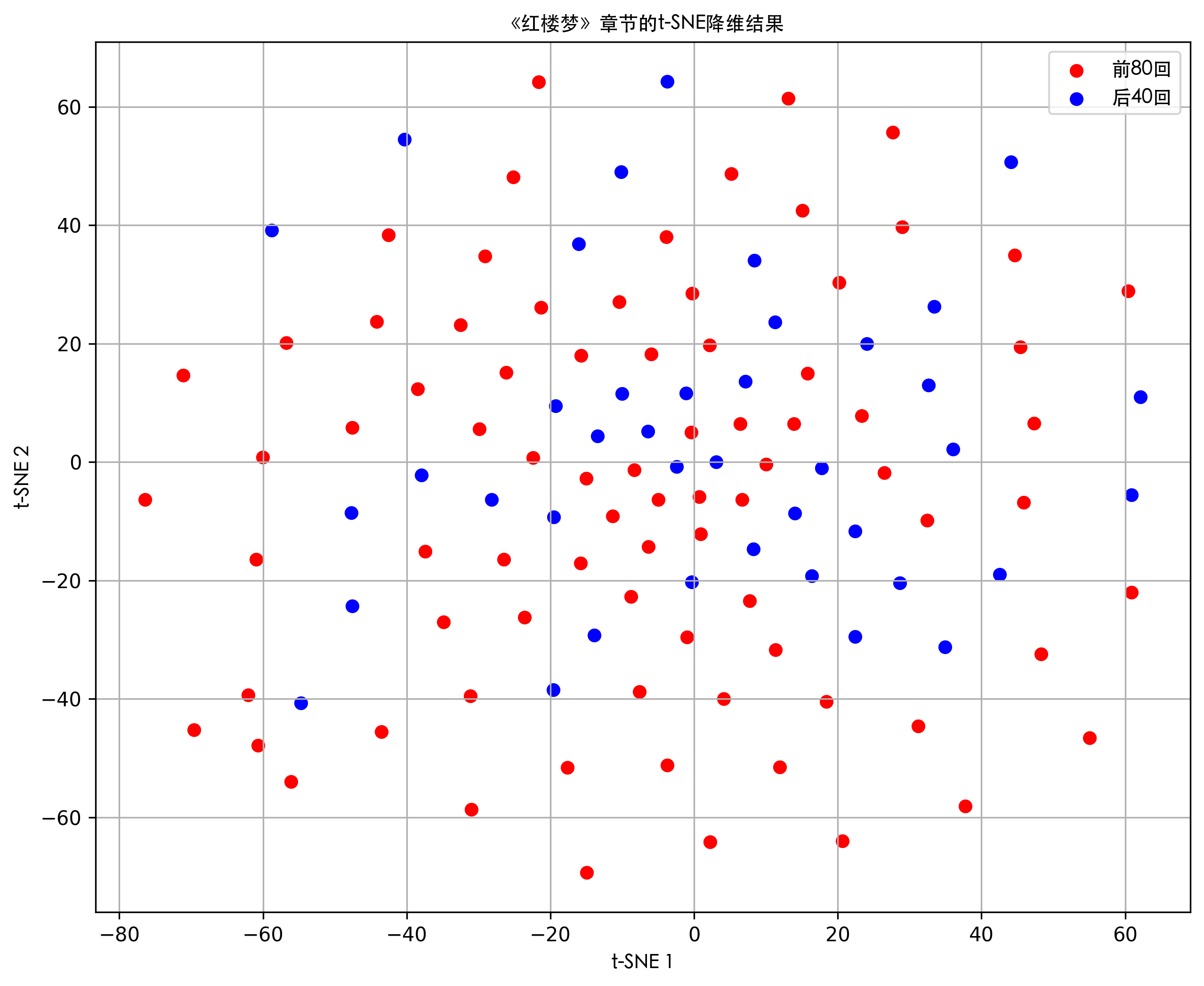
2. tfidf



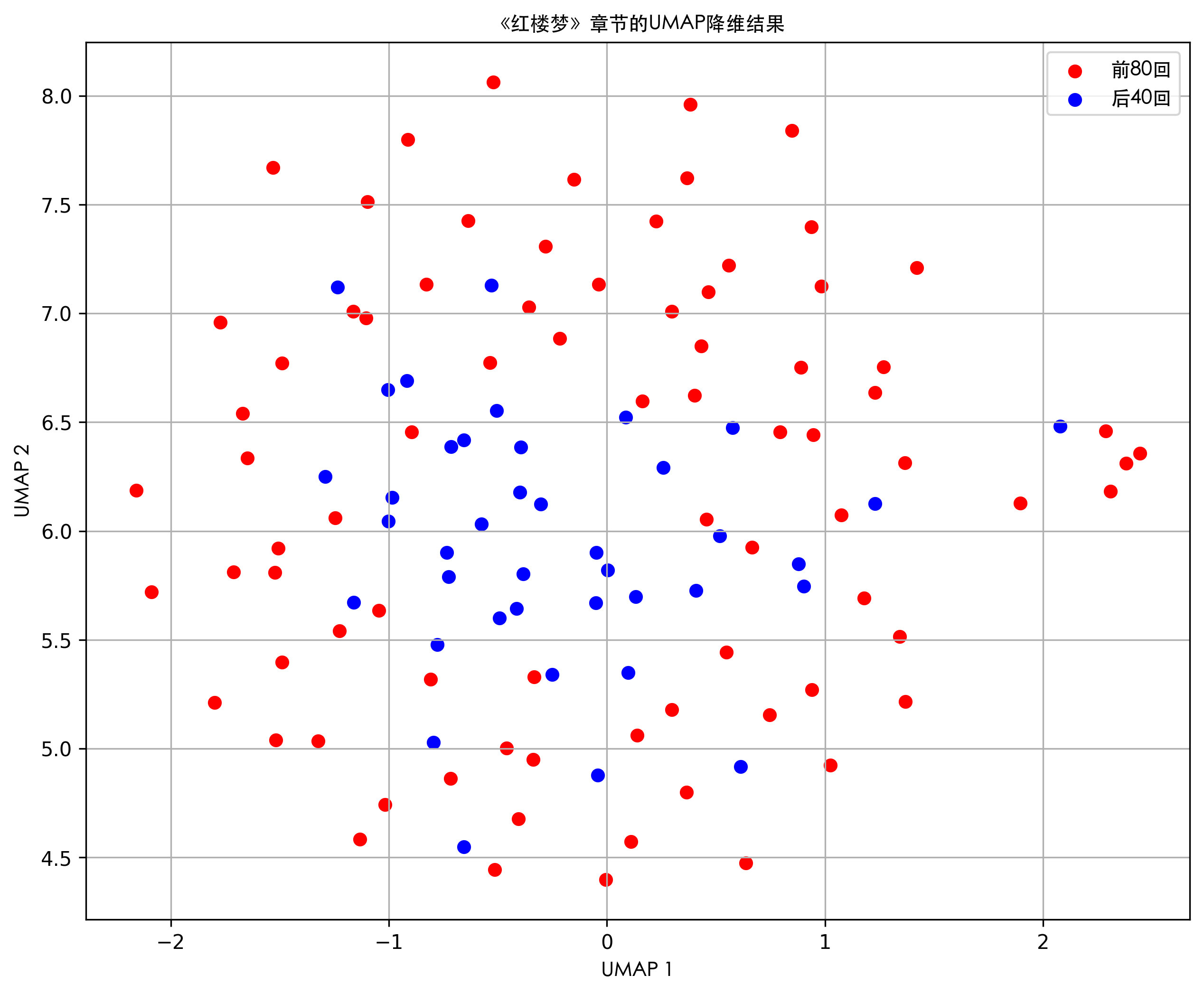
3. pca



4. tsne



5. umap



6. kmeans

