CS 498 AML Homework 4
From: Chongye Wang & Si Chen
Date: 09/29/2018

Error Under Each Category
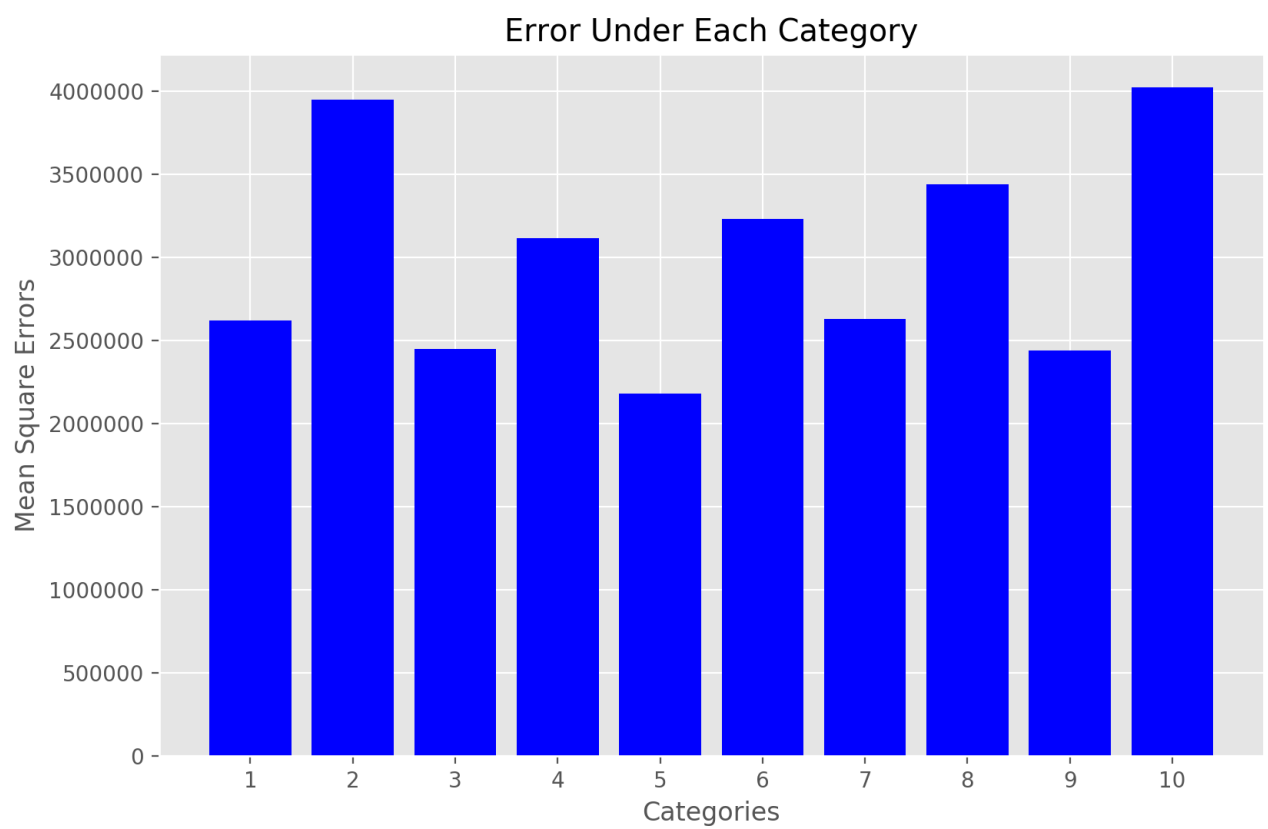
Principle Coordinate Analysis
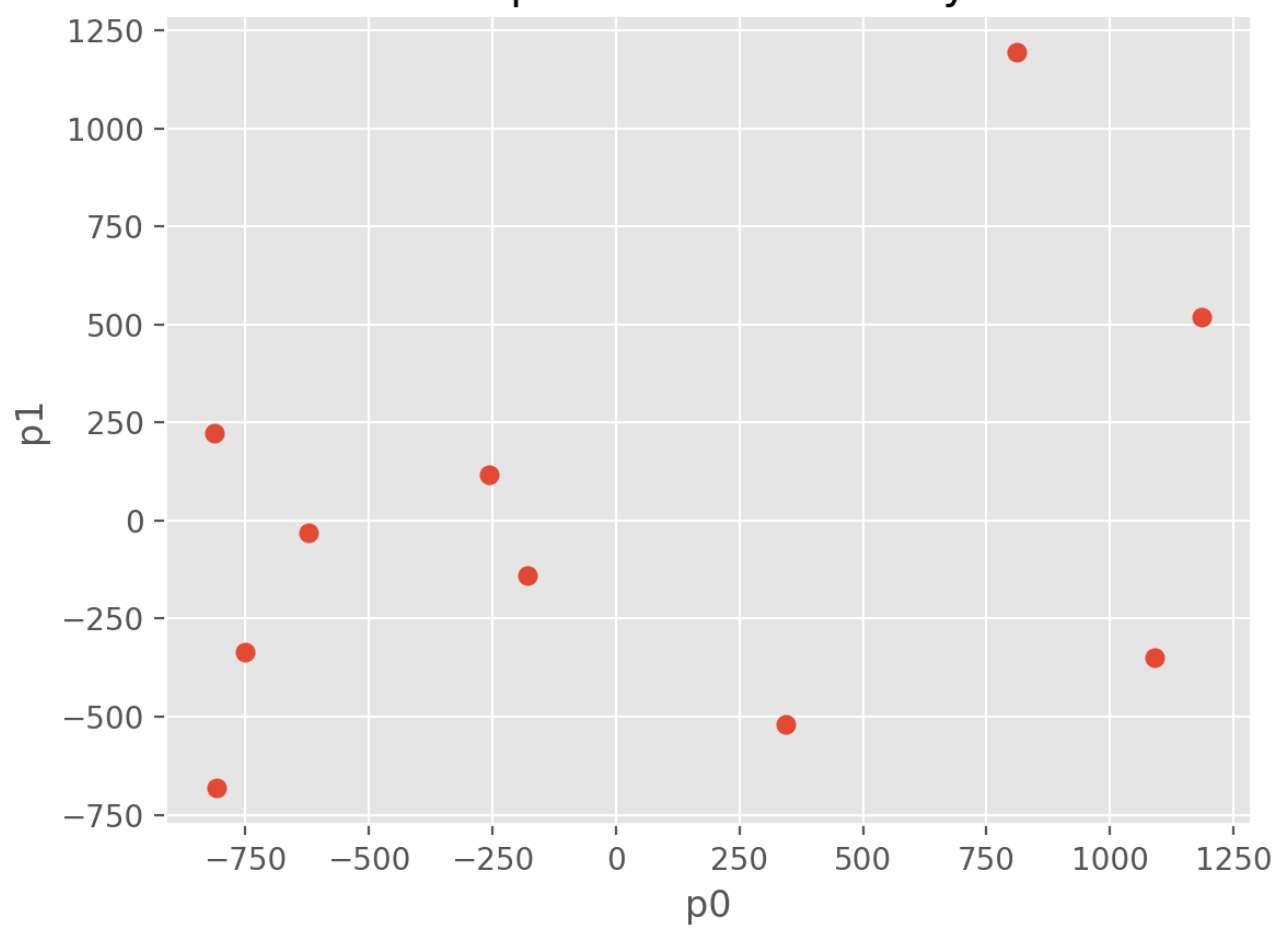
```python
error_list = []
index=[]
for label in label_data:
    sq_error=0
    data = label_data[label]
    pca = PCA(n_components = 20)
    pca_20 = pca.fit_transform(data)
    pca_reconstruction = pca.inverse_transform(pca_20)
    for i in range(len(label_data[label])):
        sq_error = sq_error + pow(np.linalg.norm(data[i] - pca_reconstruction[i]), 2)
    mse=sq_error/len(label_data[label])
    error_list.append(mse)
    index.append(label + 1)
    print(mse)

plt.style.use('ggplot')
plt.xticks(index)
plt.bar(index,error_list,fc='blue')
plt.xlabel('Categories')
plt.ylabel('Mean Square Errors')
plt.title('Error Under Each Category')
plt.show()

distance = np.zeros((10, 10))
for i in range(0, 10):
    for j in range(0, 10):
        curr1 = label_mean[i]
        curr2 = label_mean[j]
        distance[i][j] = euclidean(curr1, curr2)


mds = manifold.MDS(n_components=2, dissimilarity='precomputed', random_state=0)
pcoa = mds.fit(distance).embedding_

x=[p[0] for p in pcoa]
y=[p[1] for p in pcoa]
plt.scatter(x,y)
plt.xlabel('p0')
plt.ylabel('p1')
plt.title('Principle Coordinate Analysis')
plt.show()
```