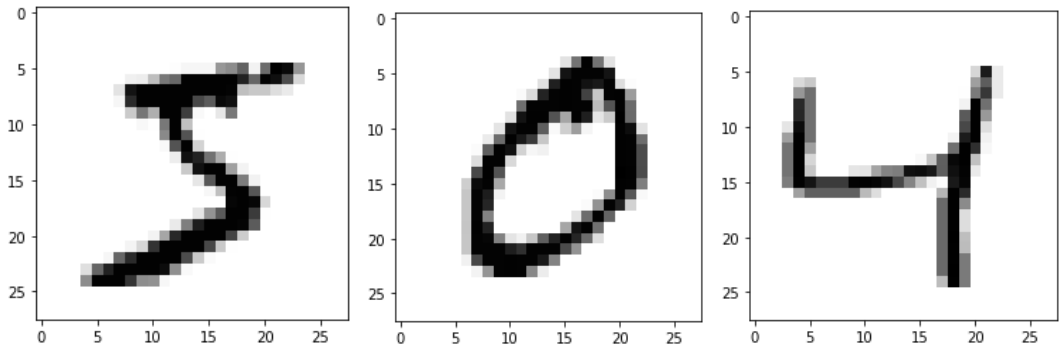
Question 1.

MNIST Data Set Samples:



Global Mean:

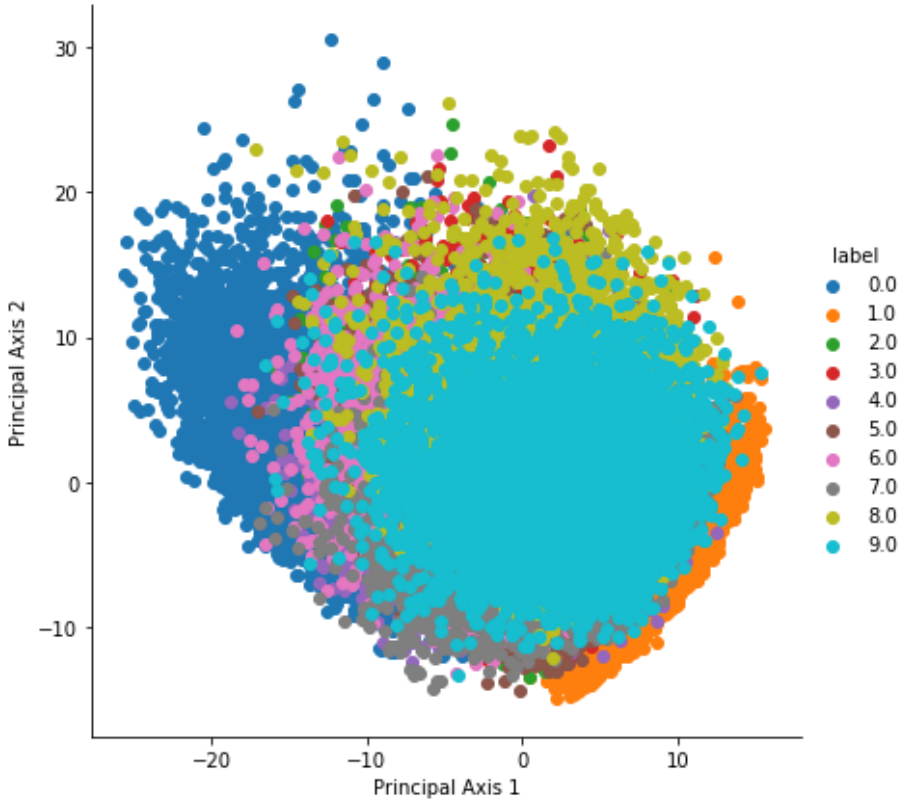
Values of 784 attributes are individually averaged out to compute the global mean.

```
Mean of the data:  [0.00000000e+00 0.00000000e+00 0.00000000e+
 0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00
 0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00
 8.23529412e-06 3.07189542e-05 1.41176471e-05 5.88235294e-07
 0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00
 0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00
 0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00
 0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00
 1.04575163e-06 3.59477124e-06 3.64052288e-05 9.52287582e-05
 1.71437908e-04 2.51372549e-04 4.71111111e-04 6.30326797e-04
 6.83071895e-04 6.95816993e-04 7.42418301e-04 6.82941176e-04
 7.33071895e-04 6.02549020e-04 3.92614379e-04 2.79346405e-04
 2.11045752e-04 8.37908497e-05 3.95424837e-05 1.38562092e-05
 0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00
 0.00000000e+00 0.00000000e+00 4.18300654e-06 2.74509804e-06
 2.72549020e-05 2.15032680e-05 1.84705882e-04 5.42745098e-04
 1.03601307e-03 1.98673203e-03 3.39921569e-03 5.05915033e-03
 7.33470588e-03 9.92137255e-03 1.25553595e-02 1.42178431e-02
 1.45960131e-02 1.33041176e-02 1.09918954e-02 8.01718954e-03
 4.71418301e-03 2.48411765e-03 1.16143791e-03 3.68562092e-04
 1.38104575e-04 3.38562092e-05 0.00000000e+00 0.00000000e+00
 0.00000000e+00 0.00000000e+00 1.26797386e-05 2.29411765e-05
 4.71241830e-05 2.73594771e-04 8.31699346e-04 2.14156863e-03
 4.52712418e-03 8.68980392e-03 1.42730719e-02 2.13254902e-02
 2.90472549e-02 3.80264052e-02 4.66003268e-02 5.19112418e-02
 5.14690850e-02 4.63275163e-02 3.74261438e-02 2.69139216e-02
 1.64456863e-02 8.92013072e-03 4.16091503e-03 1.61986928e-03
 6.35620915e-04 1.08888889e-04 1.09803922e-05 0.00000000e+00
 0.00000000e+00 2.48366013e-06 2.04575163e-05 5.62745098e-05
 3.17189542e-04 1.60869281e-03 4.09111111e-03 9.48738562e-03
 1.87284314e-02 3.29192810e-02 5.21971242e-02 7.63815686e-02
 1.06026209e-01 1.38086928e-01 1.64081569e-01 1.77461176e-01
 1.73877778e-01 1.53495882e-01 1.22855752e-01 8.98977778e-02
 5.81608497e-02 3.39662745e-02 1.78351634e-02 8.38045752e-03
 3.38392157e-03 8.14705882e-04 1.16274510e-04 7.97385621e-06
 0.00000000e+00 0.00000000e+00 4.04575163e-05 2.48431373e-04
 1.55084967e-03 5.73882353e-03 1.40697386e-02 2.83443791e-02
 5.10475163e-02 8.32075163e-02 1.23530980e-01 1.73358301e-01
 2.30995490e-01 2.89428039e-01 3.33736209e-01 3.55293268e-01
 3.48721307e-01 3.14436275e-01 2.58599281e-01 1.95269935e-01
 1.34705229e-01 8.43529412e-02 4.85894771e-02 2.60357516e-02
```

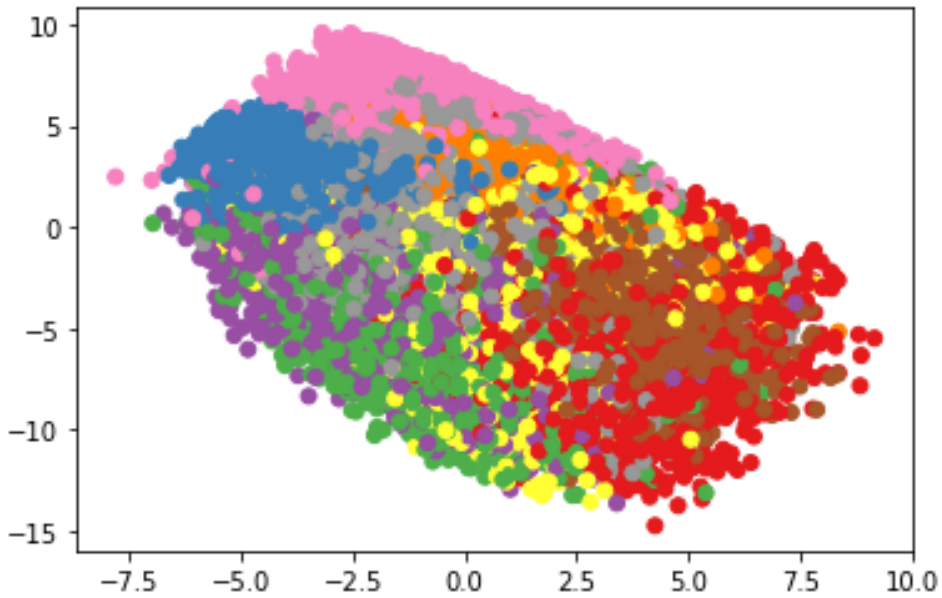
and so on.

Standard Deviation: 0.3081078038564622

PCA Data Visualization:



FDA Data Visualization:



LDA Accuracy: 88.27%

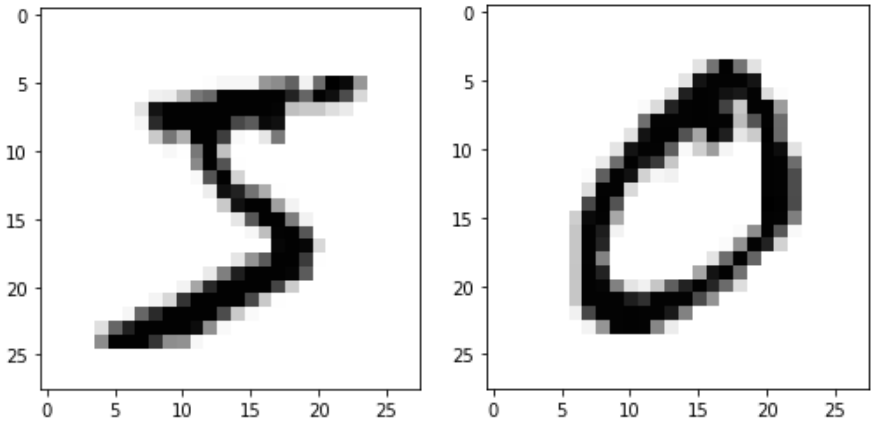
P value for:

1. Eigen Energy 95%: 179
2. Eigen Energy 70%: 24
3. Eigen Energy 90%: 132
4. Eigen Energy 99%: 322

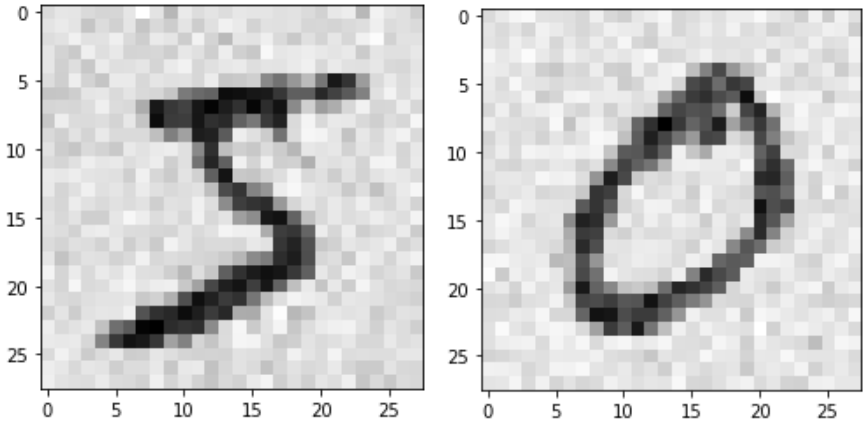
PCA, FDA and LDA have been implemented from scratch.

Question 2.

- Idx files are read to numpy arrays and normalized.
- Original data samples are shown.



- Gaussian noise is added to each pixel value.
- Images with noises are depicted.

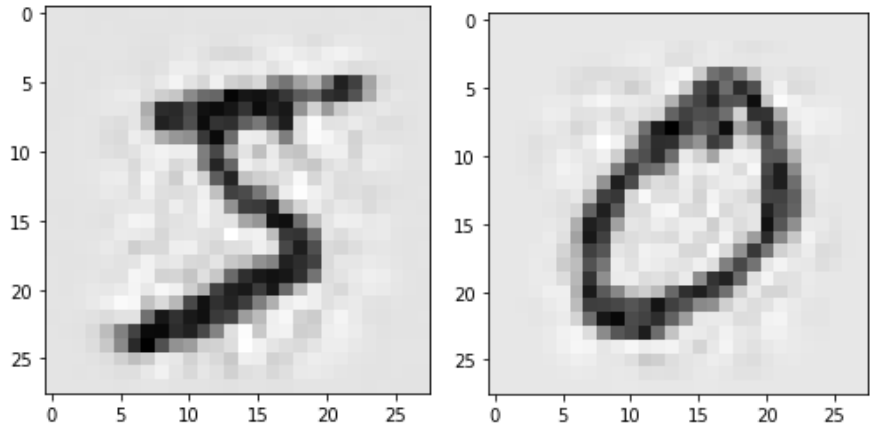


- Data is vectorized
- Calculation of mean and standard deviation.
- Covariance matrix is computed.
- Eigenvalues and eigenvectors are computed corresponding to the covariance matrix.

- Matrix multiplication of eigenvectors transform and the corrupted image data transform gives the reduced data by PCA.
- To reconstruct the images; reduced images transform and eigenvectors transform are multiplied.
- The reconstructed images are displayed.

Output on the next page.

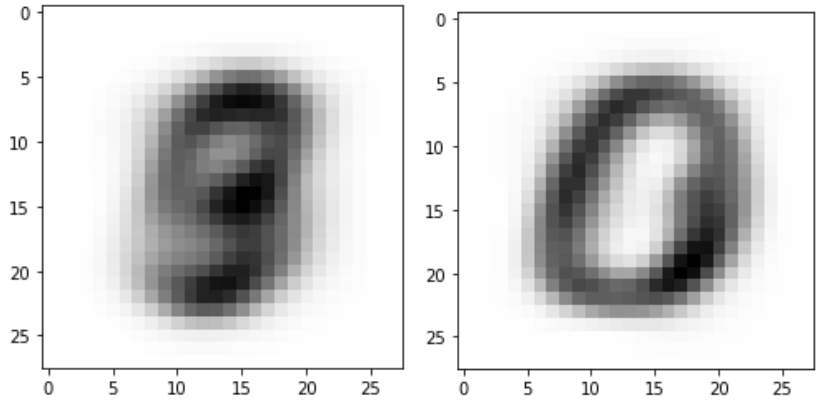
Reconstructed Output:



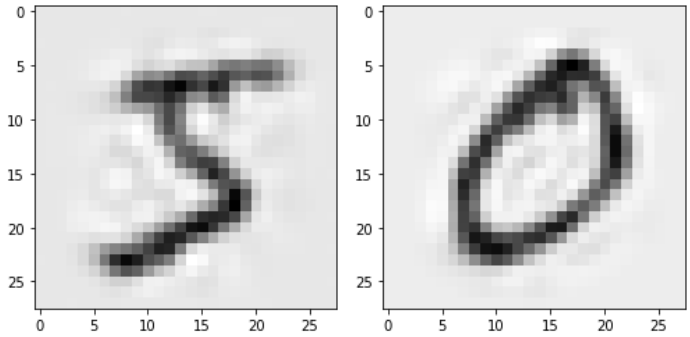
Number of components:

Observations:

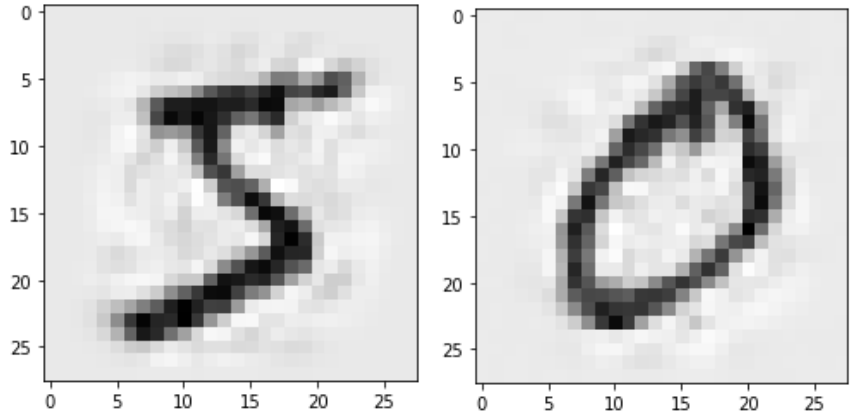
For n = 2:



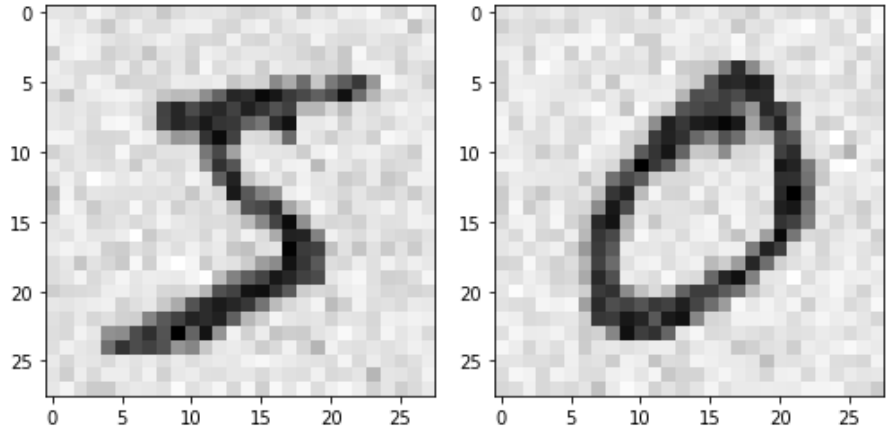
For n = 100:



For n = 200:



For n = 700:



Most noise reduction is observed around the range of n ~ (150, 200). Higher values give poor noise reduction. Lower values are hardly recognizable.