**Reid Jackson**

**4442 Assignment 2**

**February 29th 2020**

**250914839**

**1a)**

*p(Water = warm | Play = yes)* = 2/3

*p(Water = warm | Play = no)* = 1

**1b)**

*p(Play = yes | Water = warm)* = 2/3

*p(Play = no | Water = warm)* = 1/3

**1c)**

*p(Play = yes | Forecast = same)* = 1

*p(Play = yes | Forecast = change)* = 1/2

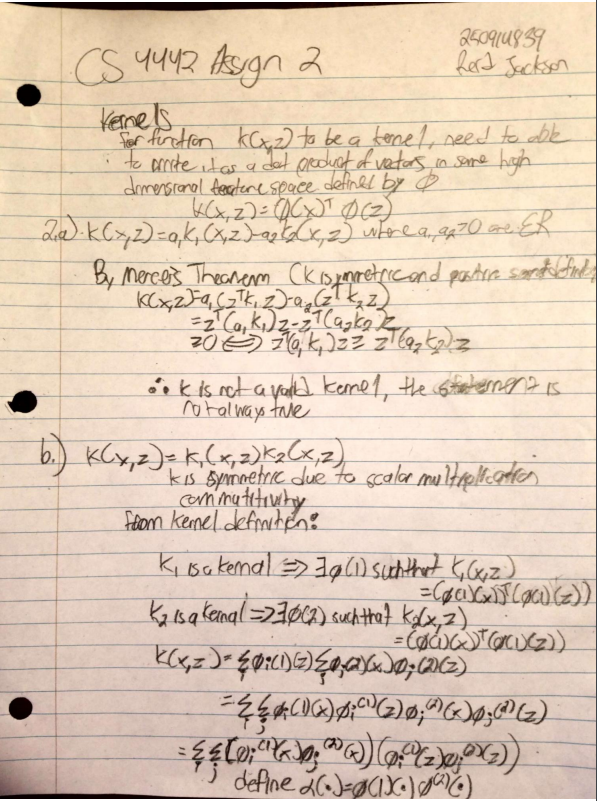
**1d)**

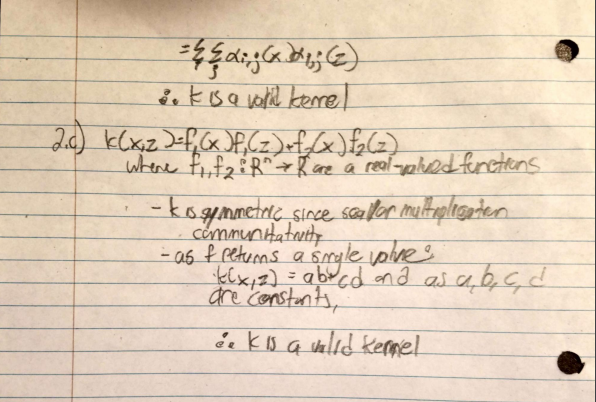
Laplace smoothing values: a = 1, d = 2

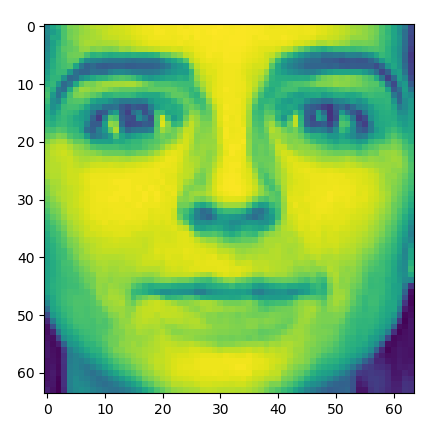
*p(Water = warm|Play = yes)* = 3/5

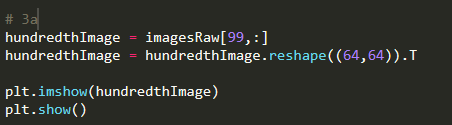
*p(Water = warm|Play = no)* = 2/3

**2)**

****

****

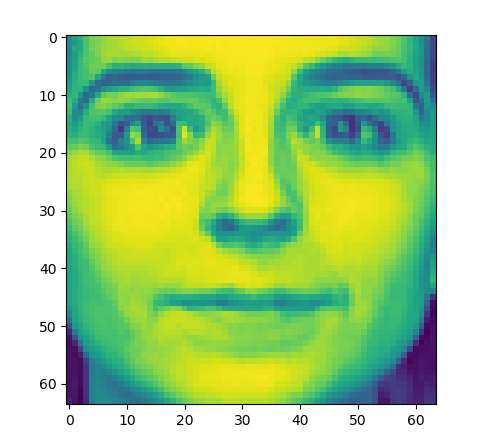
**3a)**

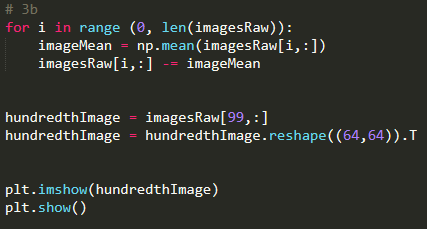
****

**3b)**

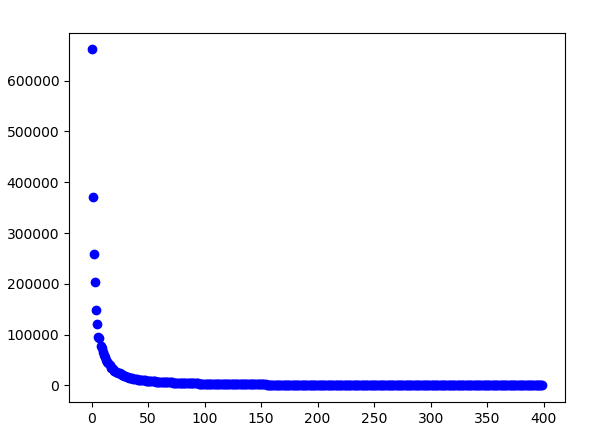
The images had a mean of: 132.38427856445313.

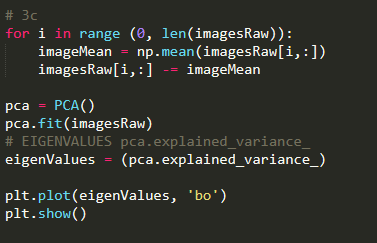
(This image is different but due to compressions the change in shading/sharpness have been lost)

****

****

**3c)**

****

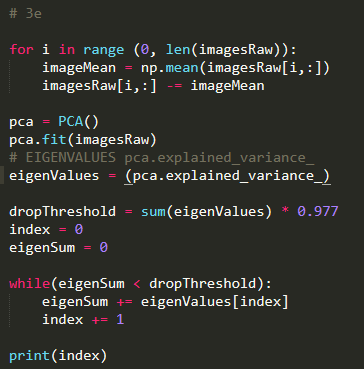
****

**3d)**

The last image, 400 reveals the least about the data after the previous 399 images have been analyzed and learned from. This eigenvalue shows that the rest of the images reveal most (99.9%+) of the variance between the images.

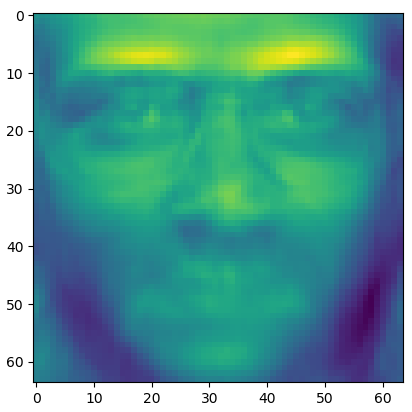
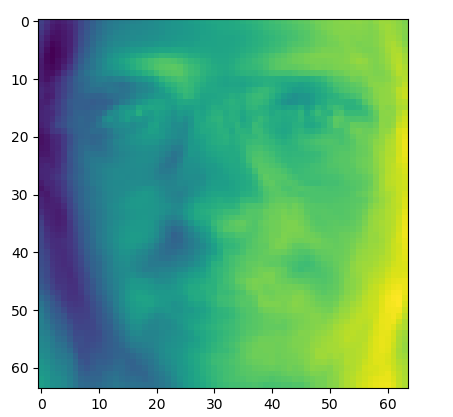
**3e)**

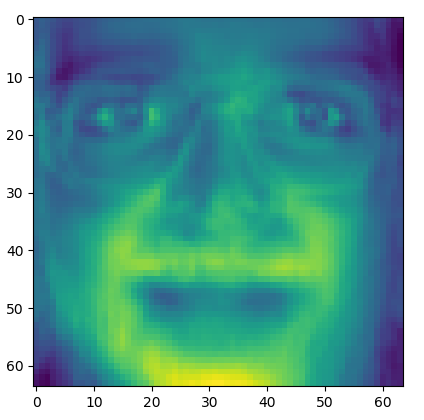
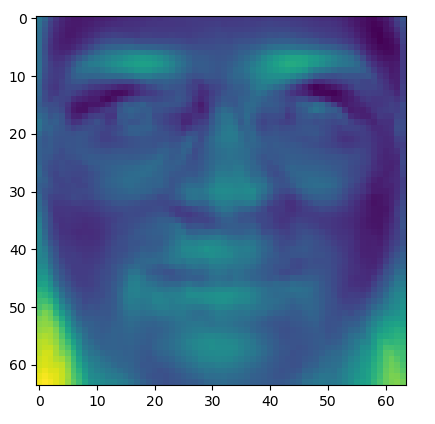
I used 97.7% for the cutoff point as this number is the value for 2 standard deviations above the mean for a stand distribution. With this cut off, *213* was the number found, so the first 213 components will be kept, while the other 187 will be dropped.

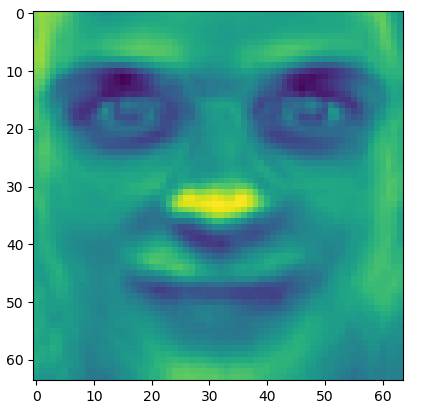


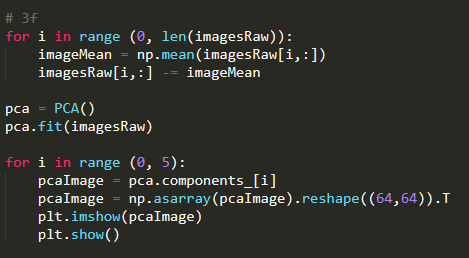
**3f)**

In order from top 1 to top 5 leading eigenvectors...



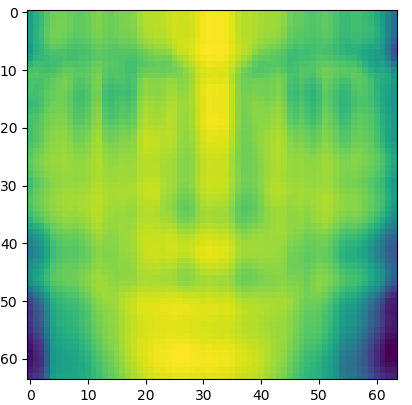
****

****

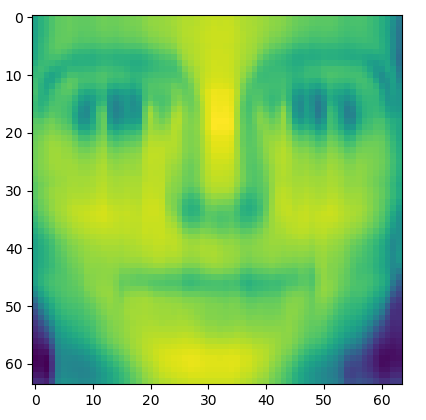
****

**3g)**

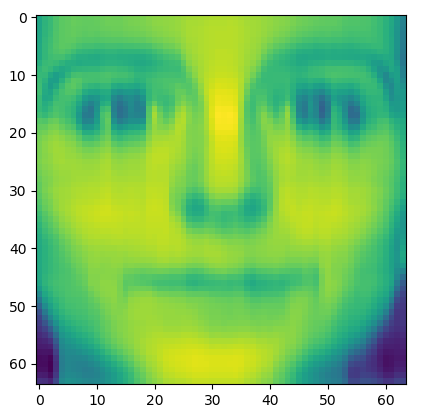
**10 Components**

****

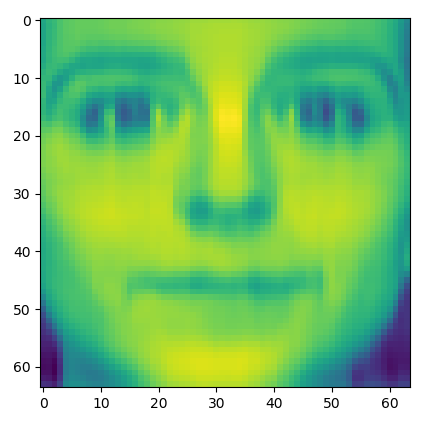
**100 Components**

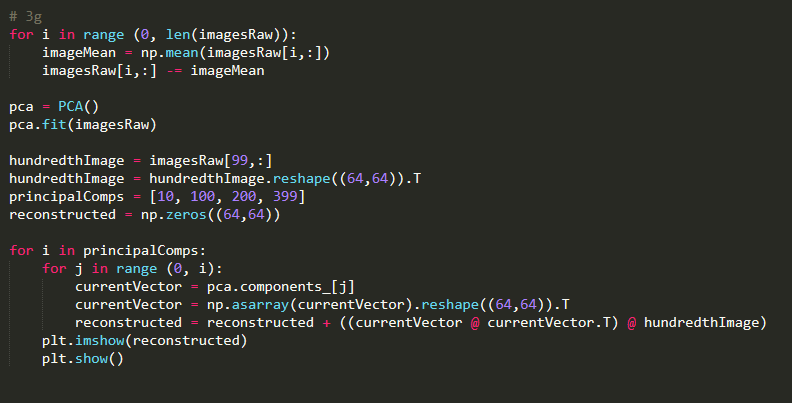
****

**200 Components**

****

**399 Components**

****

****