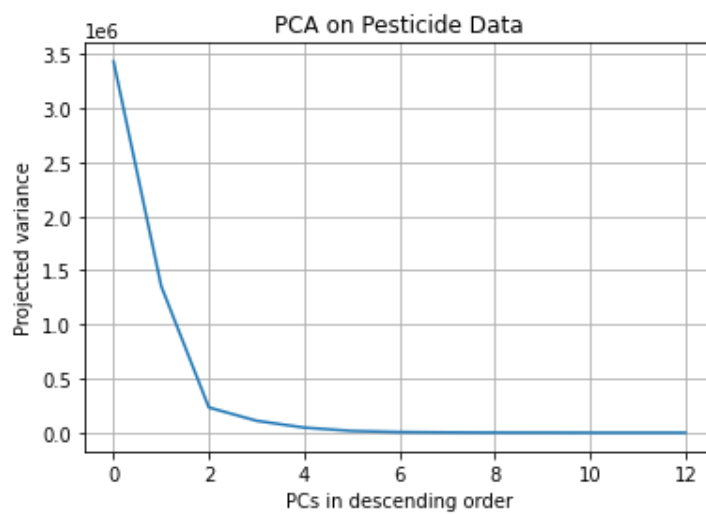
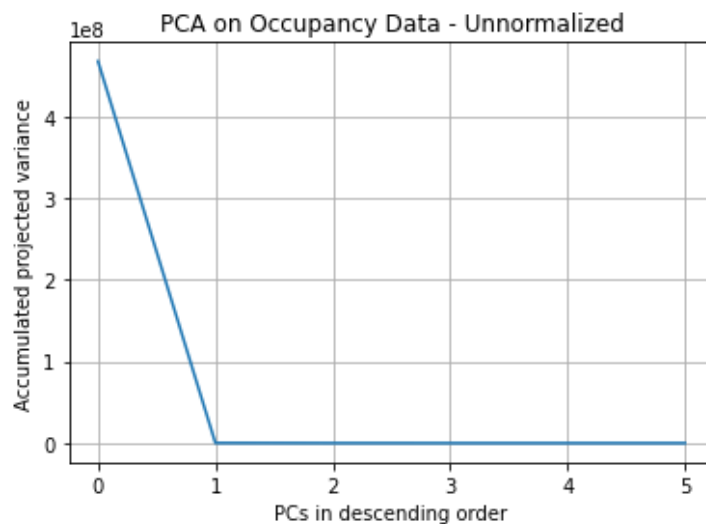
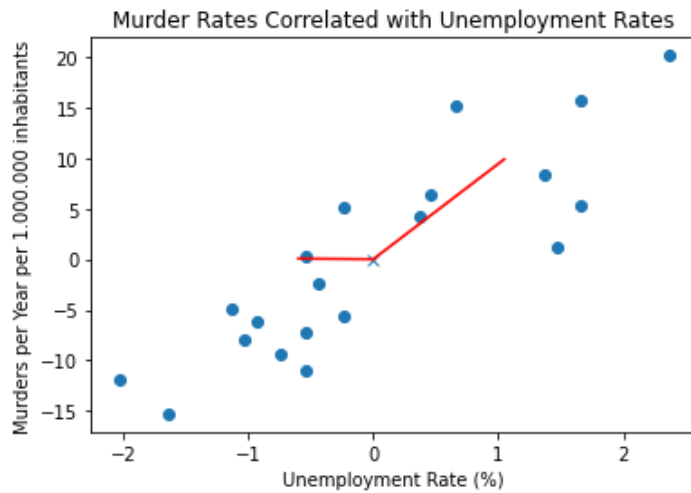
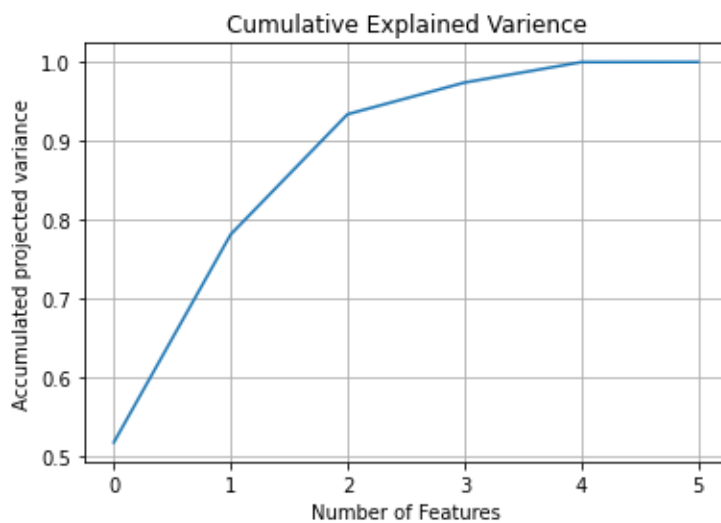


Assignment 3 – Covariance and Principal Component Analysis

Exercise 1: Performing PCA



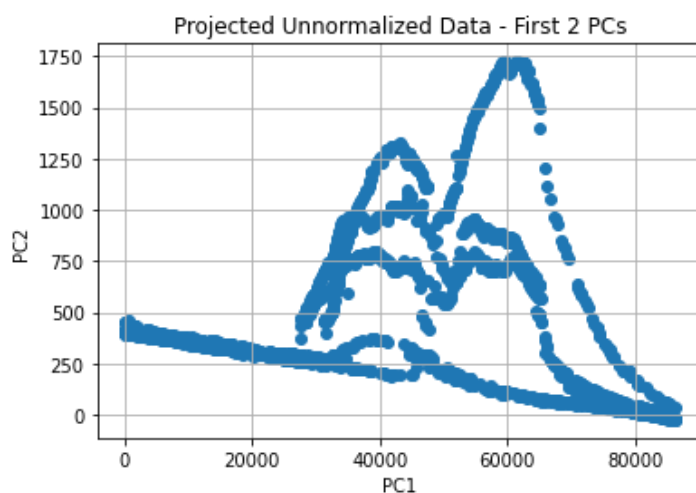
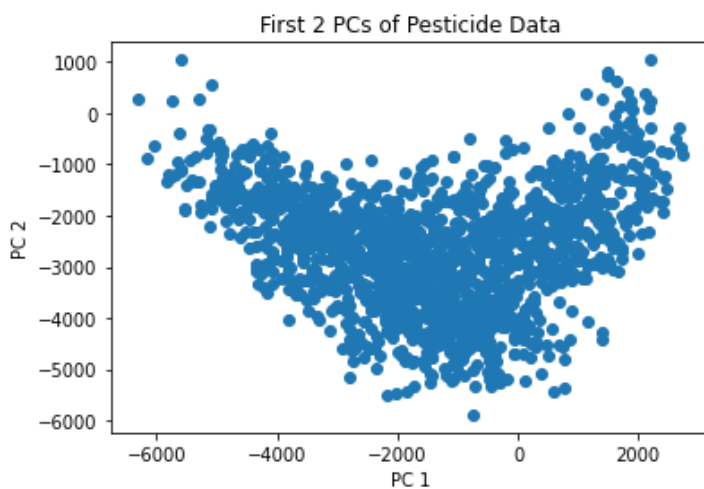


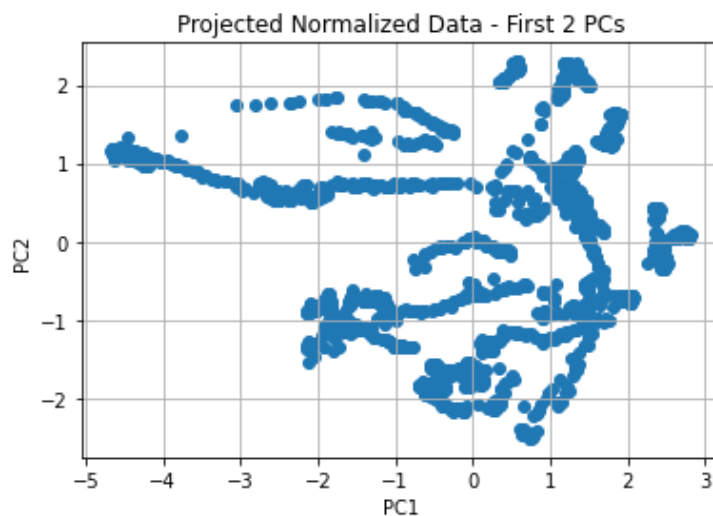
As shown in the graph and also with the following code:

```
print(c_var[0:10])  
[0.51751841 0.7813107 0.93379476 0.97393729 0.999882 1. ]
```

The number of dimensions needed to capture 90% is 3, and 4 dimensions for 95%.

Exercise 2: Visualization in 2D





Exercise 3: Clustering

To execute the code and determine the two clusters, I first stacked the XTrain and XTest data together to determine the starting point for the initialization (the first two data points). While using only the first two data points is generally not a good idea, it serves the purpose of this exercise. Then I used the KMeans function from sklearn, setting the function to determine 2 clusters, utilizing the full algorithm (as defined in the lectures), one single initialization at the starting point fitted to the XTrain data.

Cluster Center 1: `[[5.69426752e+00 4.93800425e+01 7.91594480e+02 3.84771338e+03 3.38588535e+03
1.35988535e+03 2.93734607e+02 1.31609342e+02 7.07282378e+01 3.96433121e+01 1.94437367e+01
4.23566879e+00 4.41613588e-01]]`

Cluster Center 2: `[2.19092628e+00 1.37315690e+01 1.70943289e+02 1.39436484e+03 3.18853497e+03
2.62461815e+03 1.00372023e+03 6.32814745e+02 4.95829868e+02 2.95400756e+02 1.45809074e+02
2.91984877e+01 2.83742911e+00]]`

Exercise 4: Bayesian Statistics

How is probability interpreted differently in the frequentist and Bayesian views?

The Bayesian view interprets probability as a degree of belief or a measure of certainty, while the frequentist view interprets probability as a frequency (i.e. a frequency of observing some event in a large number of trials).

Cheap, efficient computers played a major role in making Bayesian methods mainstream. Why?

Bayesian models are often analytically intractable which requires methods that are based on simulation. Cheap and efficient computers along with general purpose software allowed the models to be run more efficiently.

What is the different between a Bayesian credible interval and a frequentist confidence interval?

A Bayesian credible interval is based on a random parameter with fixed data, and results in a statement such a “I believe the value of the parameter of interest lies in that interval with a probability of ___”. While a frequentist confidence interval is based on random data and a fixed parameter, resulting in a statement such as “For the number of trials, ___% of the resulting confidence intervals will include the true value of the parameter”.

How does a maximum likelihood estimate approximate full Bayesian inference?

A maximum likelihood estimate approximates a full Bayesian inference as it uses only observed data to determine the outcome (which can be biased with small data sets), while a full Bayesian estimate incorporates prior knowledge as well as observed data to determine a more accurate conclusion.

When will point estimates be a good approximation of full Bayesian inference?

Point estimates would be a good approximation of a full Bayesian inference when there is a large data set where the posterior as Gaussian can be approximated.