# Assignment 3 - Unsupervised Dimension Reduction

Mostafa Sameer - HQ8314

**N Components and Variance:**

I chose 237 as the value for n components as the question indicated because it yielded a variance of over 90%. A comparison of various n components is shown in the graph below:

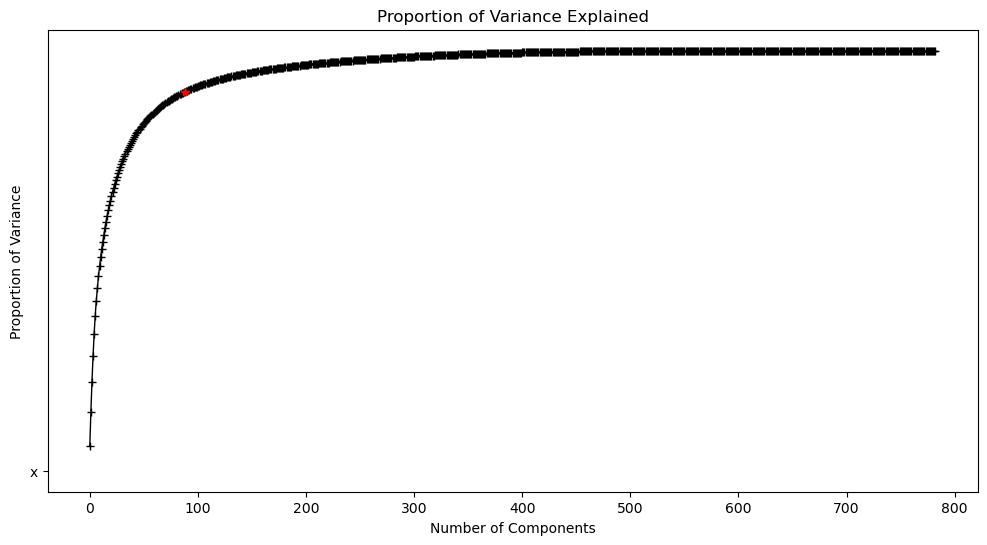


Figure: Explained Variance Proportion

**Comparison of the Classifiers:** The accuracy of the aforementioned classifiers on the original and reduced datasets are compared here:

|  |  |  |
| --- | --- | --- |
| **Classifier Name** | **Accuracy on Original Datasets** | **Accuracy on Reduced Datasets** |
| **Multinomial**  **Logistic Regression** | 76.46% | 73.2773% |
| **Naive Bayes** | 64.9765% | 23.3323% |
| **KNN** | 67.4267% | 66.3466% |