# Gene Analysis Report

## 1. Overview

This report provides a detailed analysis of the gene dataset, focusing on dimensionality reduction, clustering, and key factors contributing to gene expression variance.

## 2. Data Summary

Number of Samples: 705

Number of Features: 1713

## 3. PCA Analysis

Principal Component Analysis (PCA) was performed to reduce dimensionality.

The first two components captured a significant portion of the variance.

## 4. Clustering Analysis

K-Means clustering was applied to identify distinct groups in the gene data.

The clustering revealed three distinct groups of genes with similar expression patterns.

## 5. Gene Contributions

Total Genes: 1713

Significant Genes (contributing > 1% variance): 0

Top 5 Genes contributing most to variance:

## 6. PCA Plot with Clustering

The PCA plot with clustering is shown below.

