Principal Compontant Analysis
based on the 1000 most variant genes