2). With the help of PCA we get a line along which the dispersion is highest. Variance perpendicular to this line is minimum, so this line is the best fit line.

Comparing plot 1 and plot 2 :-

plot 1 has better approximation then plot 2. In plot 1 sum of variance perpendicular to the line is very less. where as this sum is very large in plot 2.

Also by observing the scatter plot of 2nd dataset, we see that it is similar to quadratic eq^n, so any straight line will not be a good approximation for this dataset.