# **Motivation**

Graphical user interface, text, application

Description automatically generated  
Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application, email

Description automatically generatedGraphical user interface, text, application, email

Description automatically generated

## **Data Compression**

Chart, scatter chart

Description automatically generated  
Why these points do not fall on the straight line ??  
Since the inches and cm is rounded off to their respective nearest inches and cm.

Chart

Description automatically generatedJust to summarize, if we allow ourselves to approximate the original data set by projecting all of my original examples onto this blue line in the graph, then I need only one number, I need only real number to specify the position of a point on the line, and so what I can do is therefore use just one number to represent the location of each of my training examples after they've been projected onto that green line.

Chart

Description automatically generatedWe are going to project it onto 2D. So, I've projected this data so that all of it now lies on this 2D surface. As you can see all the data lies on a plane, because we've projected everything onto a plane, and so what this means is that now I need only two numbers, z1 and z2, to represent the location of point on the plane.

Text

Description automatically generatedText, letter

Description automatically generated

Suppose if there are 50 features, then how to visualize this data 🡪 not possible

## **Visualization**

# **Principal Component Analysis PCA**

# **Applying PCA**

# **References**

<https://machinelearningmastery.com/dimensionality-reduction-for-machine-learning/>

<https://www.ritchieng.com/machine-learning-dimensionality-reduction/>