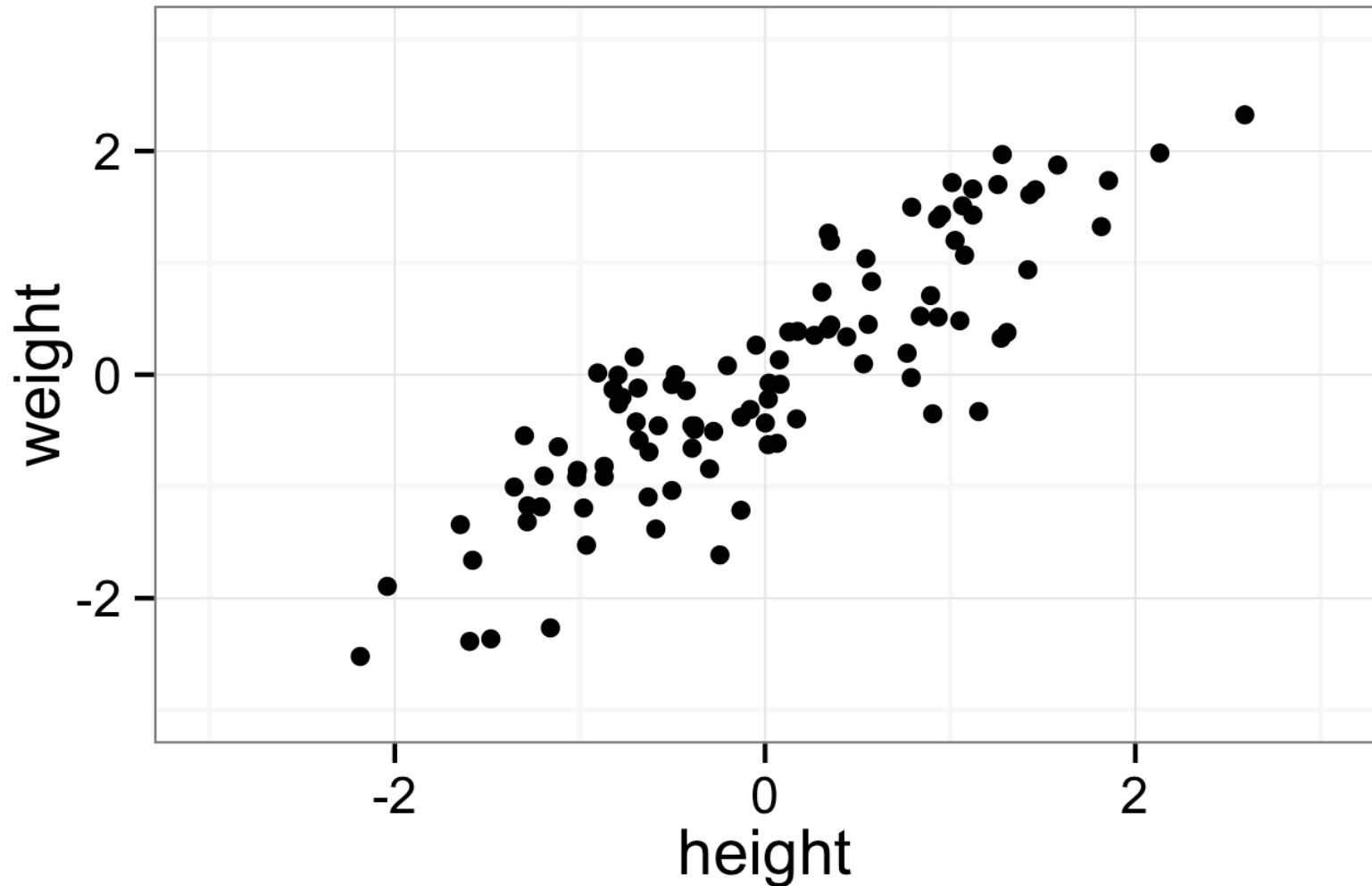


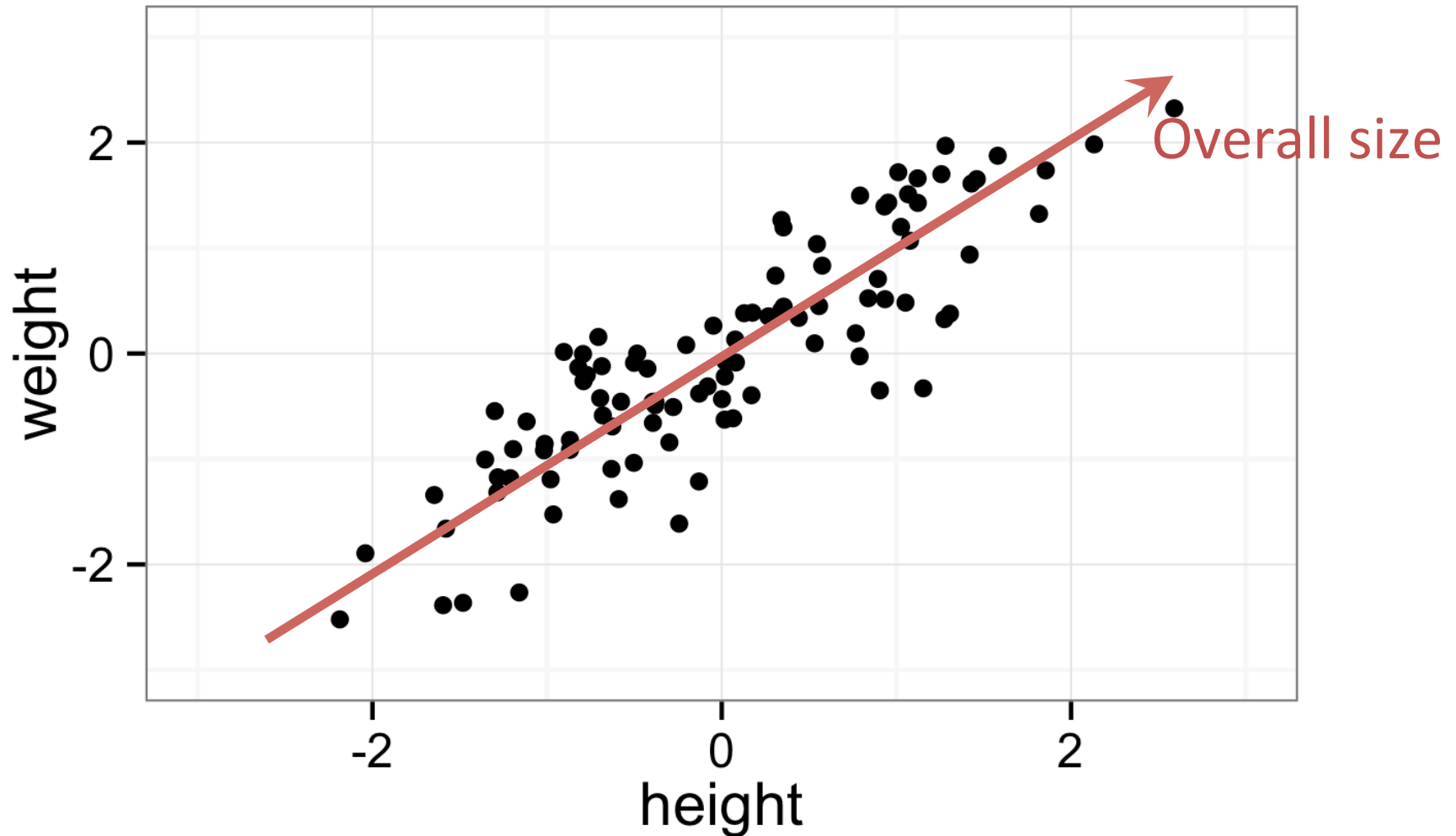
Principal Components Analysis (PCA)

Exploratory data analysis of high-dimensional data sets.

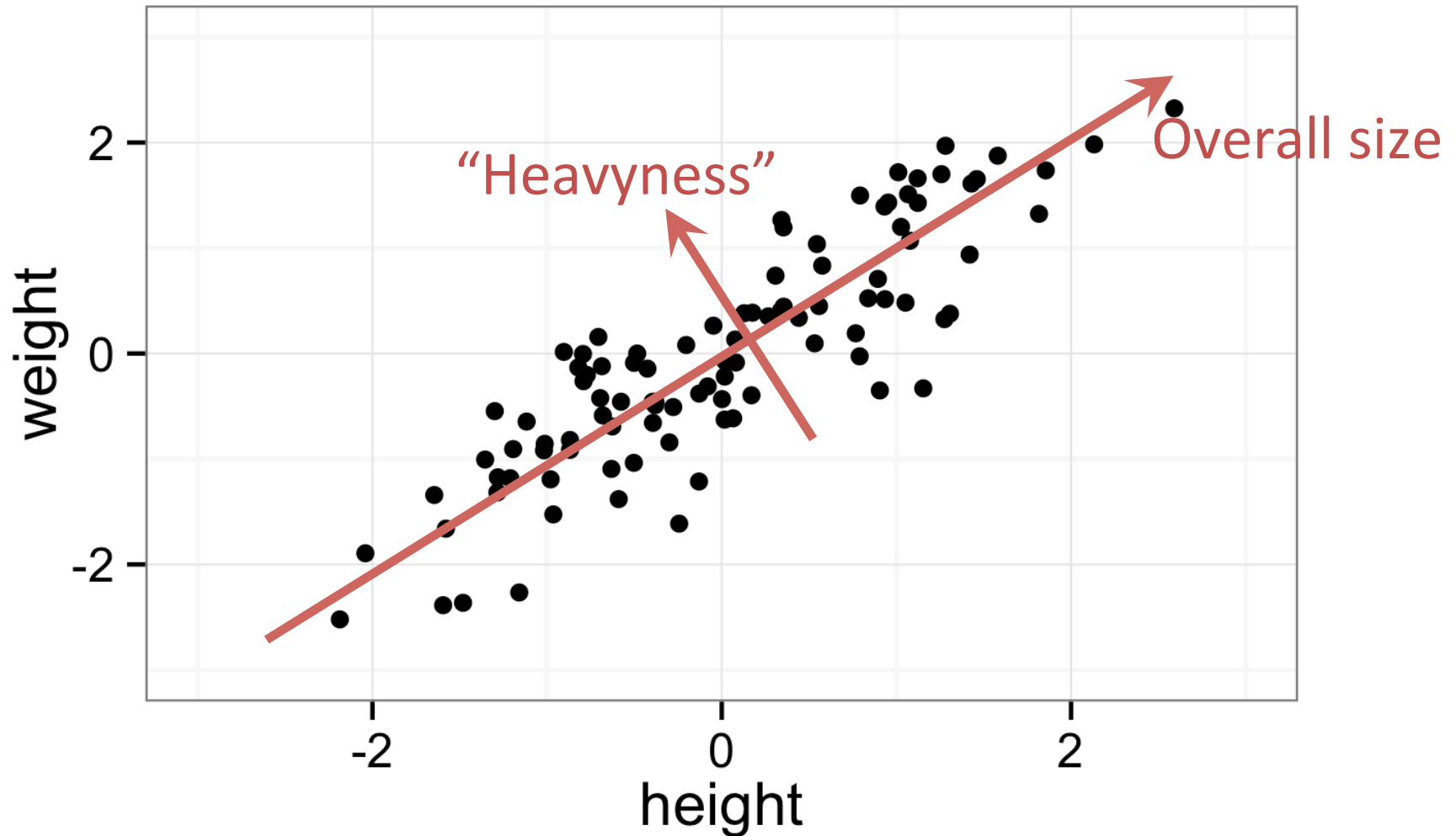
Example: Consider a data set of heights and weights of people



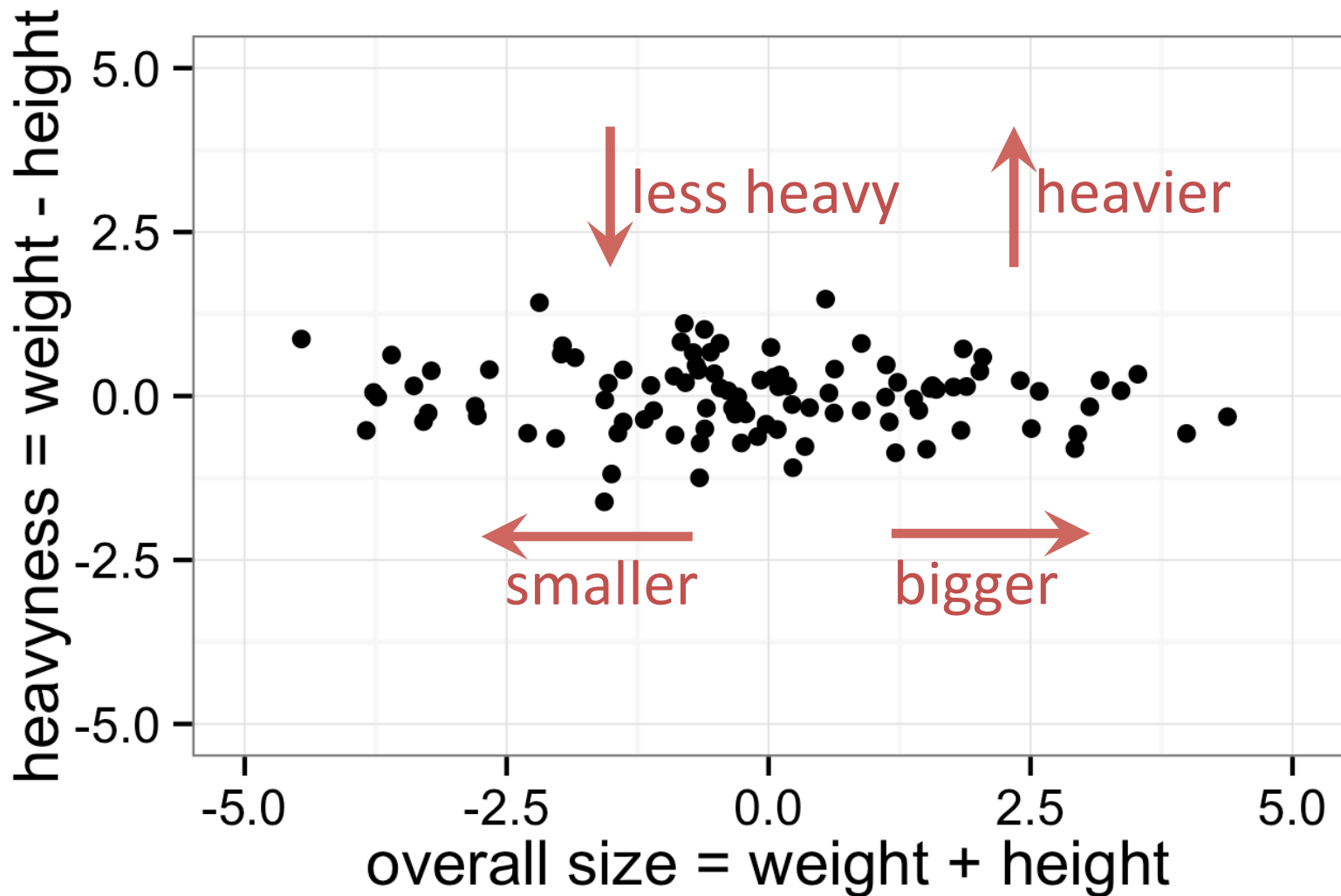
Example: Consider a data set of heights and weights of people



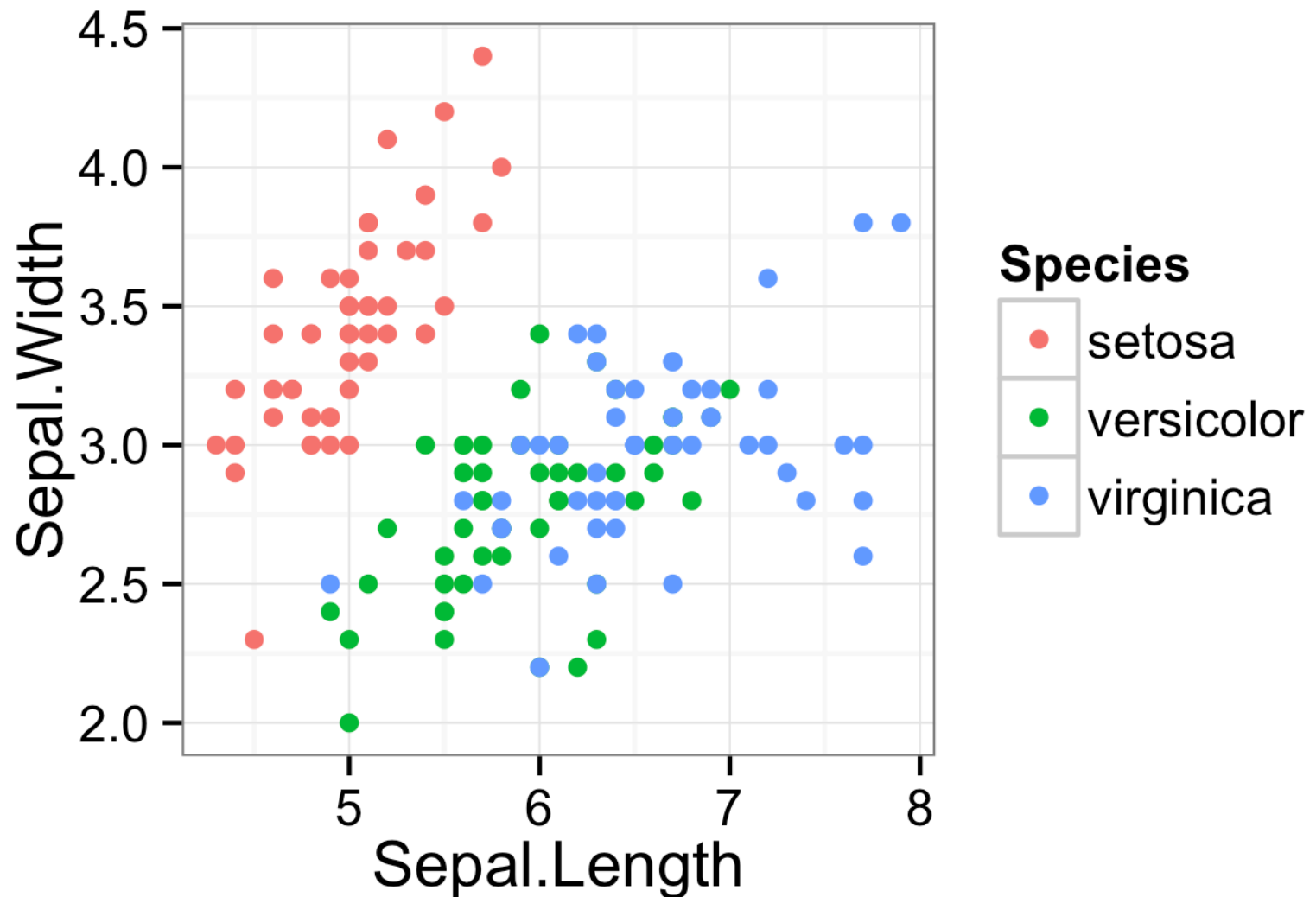
Example: Consider a data set of heights and weights of people



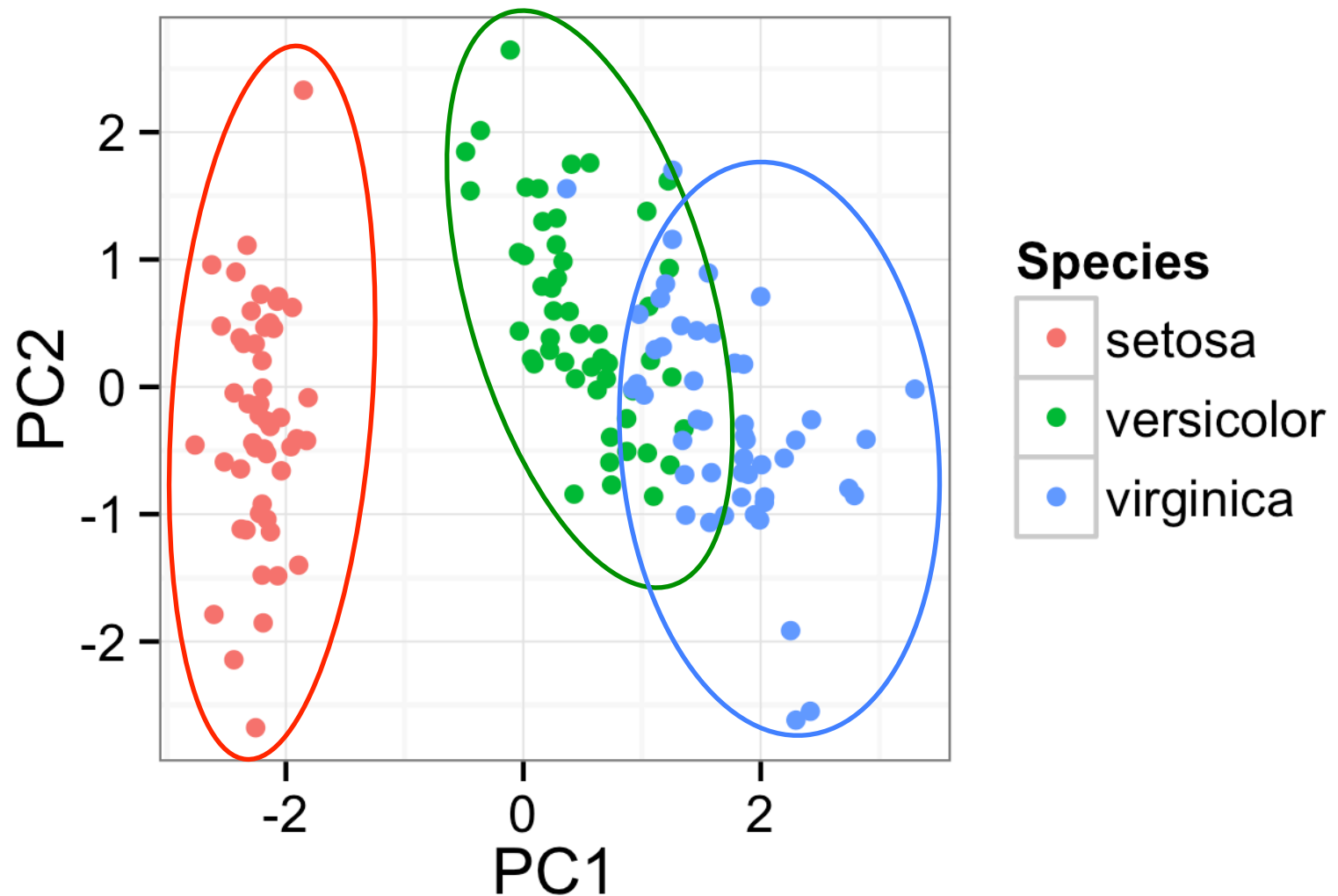
PCA on this data set reframes data in terms of overall size and fatness



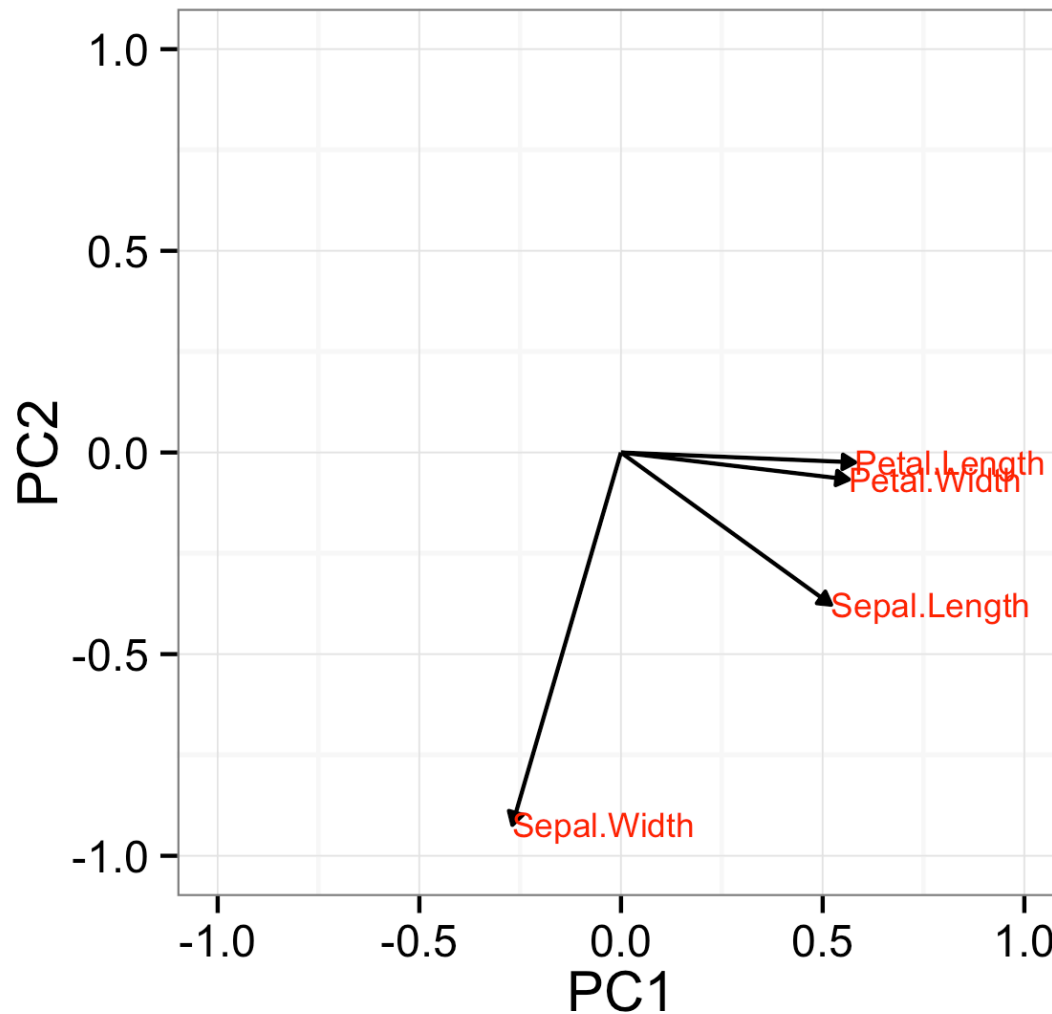
Recall the iris data set: versicolor and virginica seem to overlap



But they mostly separate when plotted as principal components (PCs)



The rotation matrix tells us which variables contribute to which PCs



The first component captures the majority of the variation in the data

