# Read training data

{ pml.training <- read.csv("~/desktop/coursera/Machine\_Learning/pml-training.csv") pml.testing <- read.csv("~/desktop/coursera/Machine\_Learning/pml-testing.csv") } # set seed and split training set{ set.seed(1000) inTrain <- createDataPartition(y=pml.training$classe, p=0.75, list=FALSE) training <- pml.training[inTrain, ] testing <- pml.training[-inTrain, ] } # data cleaning and processing ```{ dim(training) } # Trim down data set which contains 19622 observations and 160 variables