## Read the files into R variable called Activity using read.csv Note the file should be in your working directory and the code below will unzip.

unzip(zipfile="activity.zip") Activity <- read.csv("activity.csv")

## Load the ggplot2 to call the qplot function for plotting the graph.

### Use Tapply to call the function sum and remove NA's using na.rm to create a total of all the steps.

### Calculate the Mean and the median using the mean() and median() functions removing the NA's in each.

library(ggplot2) StepTotal <- tapply(Activitydate, FUN=sum, na.rm=TRUE) qplot(StepTotal, xlab="Total steps by day") mean(StepTotal, na.rm=TRUE) median(StepTotal, na.rm=TRUE)

## create a variable called intervals that uses Tapply to get the steps and the interval removing the NA's.

### use baseplot to plot the intervals with a type of 1 and titles added.

Intervals <- tapply(Activityinterval,mean, na.rm=TRUE) plot(names(Intervals), Intervals, type="l", main = "Time Series Plot", xlab="5 Min Intervals", ylab="Average Steps")

## Calculate Total number of NA's in the data set

sum(is.na(Activity))

## Based Upon total number of NA's fix the NA's in the dataset. Merge the Na's from the activity data sets with the average steps by interval from the previous graph.

SubsetNAs <- Activity[is.na(Activity),] SubsetNAsaverage\_steps

## create new data sets from the filled data above.

ActivityNoNA <- Activity ActivityNoNA[is.na(Activity),] <- SubsetNAs FinalDataSet <- tapply(ActivityNoNAdate,function(x) sum(x,na.rm=TRUE))

## Plot a histogram of number of steps with NA's removed.

hist(FinalDataSet, breaks = 25, col="red",xlab="Number of Steps", main="Daily Activity") mean(FinalDataSet, na.rm=TRUE) median(FinalDataSet, na.rm=TRUE)