NAME: . . ..

STUDENT NUMBER: . . ..

Question 1 (10 marks)

Let the sum of the squares of the decimal digits of a positive integer s_0 be represented by s_1. In a similar way, let the sum of the squares of the digits of s_1 be represented by s_2, and so on.

If $s_i = 1$ for some $i >= 0$, then the original integer s_0 is said to be **happy**.

For example, starting with 7 gives the sequence $7, 49, 97, 130, 10, 1$ so 7 is **happy**.

Write down an R function called `happy`, that takes an integer n as an input parameter and returns a boolean value indicating whether n is **happy** or not.

Indicate how you would make use of your `happy` function to generate all the **happy** numbers less than or equal to 1000.

write your question 1 script on this page

```
decimalDigits <- function(n) {
  v <- c()
  return(decimalDigitHelper(v,n))
}
```

[1]

```
decimalDigitHelper <- function(v,n) {
  if (n<10)
    return(c(v,n))
  else
    return(decimalDigitHelper(c(v,n%%10),floor(n/10)))
}
```

[2]

```
nextInteger <- function(n) {
  return(sum(decimalDigits(n)^2))
}
```

[2]

```
happy <- function(n=7) {
  s <- c(n)
  nextInt <- nextInteger(s[1])
  while( length(s[s==nextInt])==0 ) {
    s <- c(nextInt,s)
    nextInt <- nextInteger(nextInt)
  }
  return( s[1] == 1 )
}
```

[3]

```
seq <- 1:1000
seq[unlist(lapply(seq,happy))]
```

[2]

Question 2 (15 marks)

A `data.frame` called `trees` is supplied with the R distribution.

This data set provides measurements of the `Girth`, `Height` and `Volume` of timber in 31 felled black cherry trees. The girth is the diameter of the tree (in inches) measured at 4 feet 6 inches above the ground. The height is measured in feet and the volume in cubic feet.

The first few records from this data set can be examined using the `head` command:

```
> head(trees)

  Girth Height Volume
1   8.3     70   10.3
2   8.6     65   10.3
3   8.8     63   10.2
4  10.5     72   16.4
5  10.7     81   18.8
6  10.8     83   19.7
```

Measuring the volume of a felled tree is an expensive exercise and you suspect that you might be able to *model* the volume via an equation of the form:

$$myVolume = K * Girth^2 * Height$$

Construct an R script that uses the `trees` data to estimate the constant K for the model above and produce a plot that indicates how well this model does.

write your question 2 script on this page

```
t <- data.frame(GsqH = (trees$Girth ^ 2) * trees$Height,
                            Volume = trees$Volume)
```
[3]

```
# compute the slope of the best linear fit
# using least squares
(k <- sum(t$Volume*t$GsqH)/sum(t$GsqH^2))
```
[3]

```
# compute model volumes
t$myVolume <- k * t$GsqH
```
[3]

```
# and how well they correlate with measured volumes
correlation <- round(cor(t$Volume, t$myVolume),3)
```
[2]

```
# now produce a plot to show how well the model does
plot(t$Volume,t$myVolume,
              main=paste("slope of line = ",toString(1),
             "\ncorrelation = ",toString(correlation)))
abline(a=0,b=1)
```
[4]