Class06

Dhruv

General format for R script: ADD <- FUNCTION(x,y){x+y}

```
add <- function(x,y=1){x+y}
```

What would happen if we add x+y

```
add(1,1)
```

```
[1] 2
```

```
add(c(100,1,100),1)
```

```
[1] 101   2 101
```

```
add(c(100,1,100),c(100,1,100))
```

```
[1] 200   2 200
```

```
add(10)
```

```
[1] 11
```

```
add(1,1)
```

```
[1] 2
```

Make a function that generates a random nucleotide sequence of any length

```
Nucleotides <- c("A","T","G","C")
sequence <- sample(Nucleotides, 100, replace = TRUE)
sequence
```

```
  [1] "C" "A" "C" "C" "G" "C" "A" "G" "G" "C" "C" "C" "T" "G" "C" "C" "G" "G"
 [19] "T" "T" "G" "T" "G" "T" "T" "A" "T" "C" "A" "T" "A" "T" "G" "A" "C" "A"
 [37] "G" "G" "T" "C" "C" "A" "G" "A" "C" "T" "C" "A" "C" "T" "T" "A" "G" "C"
 [55] "T" "G" "C" "T" "A" "G" "A" "C" "C" "G" "T" "C" "A" "A" "A" "A" "G" "A"
 [73] "T" "C" "C" "T" "T" "C" "T" "G" "G" "C" "T" "A" "T" "T" "C" "G" "A" "G"
 [91] "G" "A" "A" "G" "C" "G" "G" "T" "T" "T"
```

```
Generate_DNA <- function(length){
  Nucleotides <- c("A","T","G","C")
  sequence <- sample(Nucleotides,size = length, replace=TRUE)
  return(sequence)
}
Generate_DNA(10)
```

```
 [1] "G" "C" "G" "T" "C" "C" "T" "A" "G" "T"
```

This working snippet ROCKS! Now i can make it into an angelic function

```
library(bio3d)
unique(bio3d::aa.table$aa1)[1:20]
```

```
 [1] "A" "R" "N" "D" "C" "Q" "E" "G" "H" "I" "L" "K" "M" "F" "P" "S" "T" "W" "Y"
[20] "V"
```

```
amino_acids <- unique(bio3d::aa.table$aa1)[1:20]
sample(amino_acids, size = 30, replace = TRUE)
```

```
 [1] "N" "T" "Q" "H" "C" "F" "P" "H" "T" "K" "H" "L" "H" "L" "T" "T" "K" "D" "H"
[20] "V" "Y" "Y" "E" "F" "L" "Y" "K" "I" "N" "F"
```

```
Generate_AA <- function(length){
  amino_acids <- unique(bio3d::aa.table$aa1)[1:20]
  string <- sample(amino_acids, size = length, replace = TRUE)
  string <- paste(string, collapse = "")
  return(string)
}
Generate_AA(100)
```

```
[1] "TIGGCTNWSTECGFHYISLVTIHQSGVTKRENQPRSVFTDERNNYNRCHGYNGYWSGTTNCVFNWFVVFEYYNWNNVWWFGSNVVIF
```

I want to generate random sequences of length 6 - 12

```
formatted<-sapply(6:12,Generate_AA)
formatted
```

```
[1] "AIAHNG"       "FTQMKLL"      "KKRQGAVH"     "LDSTFLIRH"    "CEPMIYRIRC"
[6] "FKMGYGSWKFY"  "SCPFMDLHFDFF"
```

```
cat(paste(">id.",6:12,"\n", formatted, sep = ""),sep = "\n")
```

```
>id.6
AIAHNG
>id.7
FTQMKLL
>id.8
KKRQGAVH
>id.9
LDSTFLIRH
>id.10
CEPMIYRIRC
>id.11
FKMGYGSWKFY
>id.12
SCPFMDLHFDFF
```