R Functions LabClass06

What: A function is defined with: a user selected name, a comma separated set of input arguments, and regular R code for the function body

fname<-function(arg1,arg2){paste(arg1,arg2}

My first function:

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

Can i just use it? *(specifying Y will override the y in the original function)*

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

```
[1] 2
```

```
add(x=1, y=100)
```

```
[1] 101
```

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

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

```
add(10)
```

```
[1] 11
```

You'll need to specify a third variable if you want to run this:

```
#add( 1, 1, z=1)
```

Q: make a function "generate_dna() that makes a random nucleotide sequence of any length

First, make your snippets that work

```
bases<-c("A", "C", "G", "T")
sequence<-sample(bases, size=5, replace=TRUE)
```

Then, turn it into a function

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

```
generate_dna(10)
```

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

New package! (bio3d)

There's a lot of amino acids, but we want just one entry per amino acid. Use the function `unique()` from the entire data bio3d::aa.table and use the $ to specify a column

```
unique(bio3d::aa.table$aa1)
```

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

Now let's generate a function that generates random proteins of lengths 6 to 12

```
generate_protein<-function(length){
  aa<-c(unique(bio3d::aa.table$aa1))
  protein_sequence<-sample(aa, size=length, replace=TRUE)
  #but we want to remove the " that it will give out. So use this code
  protein_sequence<-paste(protein_sequence, collapse="")
  return(protein_sequence)
}
```

```
answer<-sapply(6:12, generate_protein)
answer
```

```
[1] "VSXIMW"       "GTNIELS"      "EELYWKFP"     "NTRNGHYWL"    "AQEIIWVXAM"
[6] "KQVPYMHXXHV"  "PKPHXEYDVCRL"
```

Now let's ID each sequence:

```
paste(">id.", 6:12, answer,sep="")
```

```
[1] ">id.6VSXIMW"         ">id.7GTNIELS"        ">id.8EELYWKFP"
[4] ">id.9NTRNGHYWL"      ">id.10AQEIIWVXAM"    ">id.11KQVPYMHXXHV"
[7] ">id.12PKPHXEYDVCRL"
```

```
cat(paste(">id.", 6:12, answer,sep=""))
```

```
>id.6VSXIMW >id.7GTNIELS >id.8EELYWKFP >id.9NTRNGHYWL >id.10AQEIIWVXAM >id.11KQVPYMHXXHV >id
```

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

```
>id.6
VSXIMW >id.7
GTNIELS >id.8
EELYWKFP >id.9
NTRNGHYWL >id.10
AQEIIWVXAM >id.11
KQVPYMHXXHV >id.12
PKPHXEYDVCRL
```

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

```
>id.6
VSXIMW
>id.7
GTNIELS
>id.8
EELYWKFP
>id.9
```

```
NTRNGHYWL
>id.10
AQEIIWVXAM
>id.11
KQVPYMHXXHV
>id.12
PKPHXEYDVCRL
```