class 6

Kelly Isbell (PID: A59019188)

All about functions in R

Every function in R has at least 3 things: -name (you pick) -arguments (the input(s) to your function), and -the body

```
# Example input vectors to start with
student1 <- c(100, 100, 100, 100, 100, 100, 100, 90)
student2 <- c(100, NA, 90, 90, 90, 90, 97, 80)
student3 <- c(90, NA, NA, NA, NA, NA, NA, NA)
```

```
mean(student1)
```

```
[1] 98.75
```

```
which.min(student1)
```

```
[1] 8
```

```
mean(student1[-which.min(student1)])
```

```
[1] 100
```

```
mean(student2[-which.min(student2)])
```

```
[1] NA
```

```
  x <- student2
  mean(x[-which.min(x)], na.rm=TRUE)
```

```
[1] 92.83333
```

```
  x <- student2
  mean(x[-which.min(x)], na.rm=TRUE)
```

```
[1] 92.83333
```

```
  x <- student3
  mean(x[-which.min(x)], na.rm=TRUE)
```

```
[1] NaN
```

```
  is.na(student2)
```

```
[1] FALSE  TRUE FALSE FALSE FALSE FALSE FALSE FALSE
```

```
  x <- student3
  x[is.na(x)] <- 0
  mean(x[-which.min(x)])
```

```
[1] 12.85714
```

```
  grade <- function(x) {
    # Mask NA to zero
    x[is.na(x)] <- 0
    # Find the mean excluding the lowest score
    mean(x[-which.min(x)])
  }
```

```
  grade(student1)
```

```
[1] 100
```

```
grade(student2)
```

```
[1] 91
```

```
grade(student3)
```

```
[1] 12.85714
```

```
gradebook <- read.csv("https://tinyurl.com/gradeinput", row.names = 1)
gradebook
```

```
           hw1 hw2 hw3 hw4 hw5
student-1  100  73 100  88  79
student-2   85  64  78  89  78
student-3   83  69  77 100  77
student-4   88  NA  73 100  76
student-5   88 100  75  86  79
student-6   89  78 100  89  77
student-7   89 100  74  87 100
student-8   89 100  76  86 100
student-9   86 100  77  88  77
student-10  89  72  79  NA  76
student-11  82  66  78  84 100
student-12 100  70  75  92 100
student-13  89 100  76 100  80
student-14  85 100  77  89  76
student-15  85  65  76  89  NA
student-16  92 100  74  89  77
student-17  88  63 100  86  78
student-18  91  NA 100  87 100
student-19  91  68  75  86  79
student-20  91  68  76  88  76
```

```
grade <- function(x) {
  # Mask NA to zero
  x[is.na(x)] <- 0
  # Find the mean excluding the lowest score
  mean(x[-which.min(x)])
```

```
}

ans <- apply(gradebook, 1, grade)
ans
```

```
 student-1  student-2  student-3  student-4  student-5  student-6  student-7
     91.75      82.50      84.25      84.25      88.25      89.00      94.00
 student-8  student-9 student-10 student-11 student-12 student-13 student-14
     93.75      87.75      79.00      86.00      91.75      92.25      87.75
student-15 student-16 student-17 student-18 student-19 student-20
     78.75      89.50      88.00      94.50      82.75      82.75
```

```
#Q2
which.max(ans)
```

```
student-18
        18
```

```
#Student 18 is the top scoring student
```

```
#Q3

#Lowest score considering total scores
which.min(apply(gradebook, 2, mean, na.rm=TRUE))
```

```
hw3
  3
```

```
# Lowest score if we mask NAs to 0
mask <- gradebook
mask[is.na(mask)] <- 0
which.min(apply(mask, 2, mean))
```

```
hw2
  2
```

```r
#Q4
cor(mask$hw1, ans)
```

```
[1] 0.4250204
```

```r
cor(mask$hw2, ans)
```

```
[1] 0.176778
```

```r
cor(mask$hw3, ans)
```

```
[1] 0.3042561
```

```r
cor(mask$hw4, ans)
```

```
[1] 0.3810884
```

```r
cor(mask$hw5, ans)
```

```
[1] 0.6325982
```

```r
#HW 6 was the most predictive of overall score, it had the highest correlation.
```

```r
apply(mask, 2, cor, ans)
```

```
      hw1       hw2       hw3       hw4       hw5
0.4250204 0.1767780 0.3042561 0.3810884 0.6325982
```