

STATS380: Statistical Computing
2018 Exam
AceNighJohn

Question 1

Question 1a

```
1:3 + c(T, F, T, T, F, T)
```

```
## [1] 2 2 4 2 2 4
```

Question 1b

```
2^1:5/10
```

```
## [1] 0.2 0.3 0.4 0.5
```

Question 1c

```
matrix(1:10, 2, 5, b = TRUE)
```

```
##      [,1] [,2] [,3] [,4] [,5]  
## [1,]    1    2    3    4    5  
## [2,]    6    7    8    9   10
```

Question 1d

```
{x = c(-0.5, 1, 0.8, 1.5); pmax(0, pmin(1, x))}
```

```
## [1] 0.0 1.0 0.8 1.0
```

Question 1e

```
{x = 7; repeat {print(x); x = x + 1; if (x > 9) break}}
```

```
## [1] 7  
## [1] 8  
## [1] 9
```

Question 1f

```
levels(factor(c("foo", "boo", "loo")))[2]
```

```
## [1] "foo"
```

Question 1g

```
substring("Statistical computing", 9:12)
```

```
## [1] "cal computing" "al computing"  "l computing"   " computing"
```

Question 1h

```
{x = c(5, 1, 4, -3, 2); ifelse(x > 3, mean(x), median(x))}
```

```
## [1] 1.8 2.0 1.8 2.0 2.0
```

Question 2

Question 2a

```
seq(3, 15, by = 3)
```

Question 2b

```
10 / 2^(0:4)
```

Question 2c

```
rep(1:4, 3) + rep(0:2, each = 4)
```

Question 3

```
varcov = function(X){  
  Z = sweep(X, 2, colMeans(X))  
  (t(Z) %*% Z)/nrow(X)  
}
```

Question 4

```
pmat = function(n = 5){  
  x = 0:(n - 1)  
  outer(x, x, "+") %% n  
}
```

Question 5

Question 5a

```
plot.new()  
plot.window(xlim = c(0, 11), ylim = c(0, 4), asp = 1)  
  
cols = hcl(seq(300, 0, length = 7))  
  
xb = seq(0.5, by = 1.5, length = 7)  
xt = xb + 1  
yb = rep(1.5, 7)  
yt = yb + 1  
  
rect(xb, yb, xt, yt, col = cols, lwd = 4)  
text((xb+xt)/2, (yb+yt)/2, 1:7)  
box()
```

Question 5b

```
mat <- rbind(rep(0, 5),
             c(0, 1, 0, 4, 0),
             c(0, 2, 0, 4, 0),
             c(0, 2, 0, 5, 0),
             c(0, 3, 0, 5, 0),
             rep(0, 5))
layout(mat, heights = c(lcm(1), 1, 0.5, 0.5, 1, lcm(1)),
       widths = c(lcm(1), 1, lcm(0.5), 2, lcm(1)))
layout.show(5)
box("outer", lty = 72)
```

Question 6

Question 6a

```
text = "good and bad"
gregexpr("d", text)
```

```
## [[1]]
## [1] 4 8 12
## attr("match.length")
## [1] 1 1 1
## attr("useBytes")
## [1] TRUE
```

Question 6b

```
df1 = data.frame(n = c("a", "b", "c"), x = 1:3)
df2 = data.frame(n = c("a", "c"), y = c(6, 9))
merge(df1, df2)
```

```
##   n x y
## 1 a 1 6
## 2 c 3 9
```

Question 6c

```
breakText = function(t){
  strsplit(t, " ")[[1]]
}
text = c("roses are red", "and so are you")
lapply(text, breakText)
```

```
## [[1]]
## [1] "roses" "are"   "red"
##
## [[2]]
## [1] "and" "so"   "are" "you"
```

Question 6d

```
text1 = "John (fishing, hunting), Paul (hiking, biking),"
text2 = "Carol, Smith (fishing, swimming)"
text = paste(text1, text2)
newtext = gsub("\\(\\.[^)]*?,) ", "\\1", text)
newtext = strsplit(newtext, ", ")[[1]]
regmatches(newtext, regexpr(",]", newtext)) = ", "
newtext

## [1] "John (fishing, hunting)" "Paul (hiking, biking)"
## [3] "Carol"                  "Smith (fishing, swimming)"
```

Question 6e

```
text = c("Smith,(919)319-1677", "Ali, 800-899-2164",
        "Richard, 7042982145")
patt1 = "\\([2-9]\\d\\d\\d\\) ?[2-9]\\d\\d\\d-\\d\\d\\d\\d\\d"
patt2 = "[2-9]\\d\\d\\d-[2-9]\\d\\d\\d-\\d\\d\\d\\d\\d"
patt = paste0(c("(", patt1, ")|", "(", patt2, ")"), collapse = "")
data = do.call("rbind", strsplit(text, ","))
data = as.data.frame(data)
colnames(data) = c("name", "phone")
pos = grepl(patt, data[, 2])
data$valid = ifelse(pos, "valid", "invalid")
data

##      name      phone  valid
## 1 Smith (919)319-1677  valid
## 2 Ali 800-899-2164    valid
## 3 Richard 7042982145 invalid
```

Question 7

```
pop.mean.df = aggregate(population ~ country, data = data, FUN = mean)
colnames(pop.mean.df) = c("country", "pop.mean")
combined.df = merge(data, pop.mean.df, by = "country")
subset(combined.df, population > pop.mean)
```

Question 8

Question 8a

```
SalesLong <- melt(Sales, id = c("Year", "Month"))
```

Question 8b

```
result <- aggregate(value ~ Month, data = SalesLong, FUN = sum)
```

Question 8c

```
small = min(result$value)
subset(result, small == value)
```