Bios 6301: Assignment 8

Yiqing Pan

30

Due Tuesday, 12 November, 1:00 PM

$5^{n=day}$ points taken off for each day late.

30 points total.

Submit a single quarto file (named `homework8.qmd`), along with a valid PDF output file. Inside the file, clearly indicate which parts of your responses go with which problems (you may use the original homework document as a template). Add your name as `author` to the file's metadata section. Raw R code/output or word processor files are not acceptable.

Failure to name file `homework8.qmd` or include author name may result in 5 points taken off.

Question 1

15 points

Install the `readxl` package and run the following

```
fn <- 'icd10.xlsx'
if(file.access(fn, mode = 4) == -1) {
    url <- "https://www.cdc.gov/nhsn/xls/icd10-pcs-pcm-nhsn-opc.xlsx"
    download.file(url, destfile = fn, mode = 'wb')
}
dat <- readxl::read_excel(fn, sheet = 2)
```

1. Show the class of `dat`. (1 point)

```
class(dat)
```

```
## [1] "tbl_df"     "tbl"        "data.frame"
```

2. Show the methods available for objects of the given class (if there are multiple classes, show methods for all classes). (3 points)

```
methods(class = "tbl_df")
```

```
##  [1] [             [[            [[<-          [<-           $
##  [6] $<-           as.data.frame coerce        initialize    names<-
## [11] Ops           row.names<-   show          slotsFromS3   str
## see '?methods' for accessing help and source code
```

```
methods(class = "tbl")
```

```
##  [1] [[<-        [<-         $<-         coerce      format      initialize
##  [7] Ops         print       show        slotsFromS3
## see '?methods' for accessing help and source code
```

```
methods(class = "data.frame")
```

```
##  [1] [             [[            [[<-          [<-           $<-
##  [6] aggregate     anyDuplicated anyNA         as.data.frame as.list
## [11] as.matrix     by            cbind         coerce        dim
## [16] dimnames      dimnames<-    droplevels    duplicated    edit
## [21] format        formula       head          initialize    is.na
## [26] Math          merge         na.exclude    na.omit       Ops
## [31] plot          print         prompt        rbind         row.names
## [36] row.names<-   rowsum        show          slotsFromS3   split
## [41] split<-       stack         str           subset        summary
## [46] Summary       t             tail          transform     type.convert
## [51] unique        unstack       within        xtfrm
## see '?methods' for accessing help and source code
```

3. If you call `print(dat)`, what print method is being dispatched? (1 point)

```
# print(dat)
# methods("print")
```

Since dat is in the tibble class,it uses print.tbl method.

4. Set the class of `dat` to be a data.frame. (1 point)

```
class(dat) <- "data.frame"
```

5. If you call `print(dat)` again, what print method is being dispatched? (1 point)

```
#print(dat)
```

After changing dat to a data.frame, calling print(dat) would dispatch the print.data.frame method.

Define a new generic function `nUnique` with the code below.

```
nUnique <- function(x) {
    UseMethod('nUnique')
}
```

6. Write a default method for `nUnique` to count the number of unique values in an element. (2 points)

```
nUnique.default <- function(x) {
    length(unique(x))
}
```

7. Check your function (2 points)

```r
nUnique(letters) # should return 26
nUnique(sample(10, 100, replace = TRUE)) # should return 10 (probably)
```

8. Write a data.frame method for `nUnique` to operate on data.frame objects. This version should return counts for each column in a data.frame. (2 points)

```r
nUnique.data.frame <- function(x) {
    sapply(x, function(col) length(unique(col)))
}
```

9. Check your function (2 points)

```r
nUnique(dat)
```

Question 2

15 points

Programming with classes. The following function will generate random patient information.

```r
makePatient <- function() {
  vowel <- grep("[aeiou]", letters)
  cons <- grep("[^aeiou]", letters)
  name <- paste(sample(LETTERS[cons], 1), sample(letters[vowel], 1), sample(letters[cons], 1), sep='')
  gender <- factor(sample(0:1, 1), levels=0:1, labels=c('female','male'))
  dob <- as.Date(sample(7500, 1), origin="1970-01-01")
  n <- sample(6, 1)
  doa <- as.Date(sample(1500, n), origin="2010-01-01")
  pulse <- round(rnorm(n, 80, 10))
  temp <- round(rnorm(n, 98.4, 0.3), 2)
  fluid <- round(runif(n), 2)
  list(name, gender, dob, doa, pulse, temp, fluid)
}
```

1. Create an S3 class `medicalRecord` for objects that are a list with the named elements `name`, `gender`, `date_of_birth`, `date_of_admission`, `pulse`, `temperature`, `fluid_intake`. Note that an individual patient may have multiple measurements for some measurements. Set the RNG seed to `8` and create a medical record by taking the output of `makePatient`. Print the medical record, and print the class of the medical record. (5 points)

```r
set.seed(8)

medicalRecord <- function() {
  vowel <- grep("[aeiou]", letters)
  cons <- grep("[^aeiou]", letters)
  name <- paste(sample(LETTERS[cons], 1), sample(letters[vowel], 1), sample(letters[cons], 1), sep='')
  gender <- factor(sample(0:1, 1), levels=0:1, labels=c('female','male'))
  dob <- as.Date(sample(7500, 1), origin="1970-01-01")
  n <- sample(6, 1)
  doa <- as.Date(sample(1500, n), origin="2010-01-01")
  pulse <- round(rnorm(n, 80, 10))
```

```
  temp <- round(rnorm(n, 98.4, 0.3), 2)
  fluid <- round(runif(n), 2)
  list(name = name, gender = gender, dob = dob, doa = doa, pulse = pulse, temp = temp, fluid = fluid)
}


patient <- medicalRecord()
class(patient) <- "medicalRecord"

print(patient)
```

```
## $name
## [1] "Yes"
##
## $gender
## [1] male
## Levels: female male
##
## $dob
## [1] "1977-05-03"
##
## $doa
## [1] "2013-06-09" "2013-07-02"
##
## $pulse
## [1] 79 78
##
## $temp
## [1] 98.07 97.50
##
## $fluid
## [1] 0.28 0.52
##
## attr(,"class")
## [1] "medicalRecord"
```

```
print(class(patient))
```

```
## [1] "medicalRecord"
```

2. Write a `medicalRecord` method for the generic function `mean`, which returns averages for pulse, temperature and fluids. Also write a `medicalRecord` method for `print`, which employs some nice formatting, perhaps arranging measurements by date, and `plot`, that generates a composite plot of measurements over time. Call each function for the medical record created in part 1. (5 points)

```
medicalRecord <- function(x) {
  UseMethod("medicalRecord")
}


medicalRecord.mean <- function(x) {
  return(list(
```

```
    mean_pulse = mean(x$pulse),
    mean_temperature = mean(x$temp),
    mean_fluid_intake = mean(x$fluid)
  ))
}

medicalRecord.print <- function(x) {
  cat("Name:", x$name, "\n")
  cat("Gender:", as.character(x$gender), "\n")
  cat("Date of Birth:", format(x$dob, "%Y-%m-%d"), "\n\n")

  cat("Test Results:\n")
  admission_data <- data.frame(
    date_of_admission = x$doa,
    pulse = x$pulse,
    temperature = x$temp,
    fluid_intake = x$fluid
  )

  print(admission_data)
}

medicalRecord.plot <- function(x) {
  dates <- x$doa
  par(mfrow = c(3, 1))

  plot(dates, x$pulse, type = 'b', xlab = "Date of Admission", ylab = "Pulse")
  plot(dates, x$temperature, type = 'b', xlab = "Date of Admission", ylab = "Temperature")
  plot(dates, x$fluid_intake, type = 'b', xlab = "Date of Admission", ylab = "Fluid Intake")
}
```

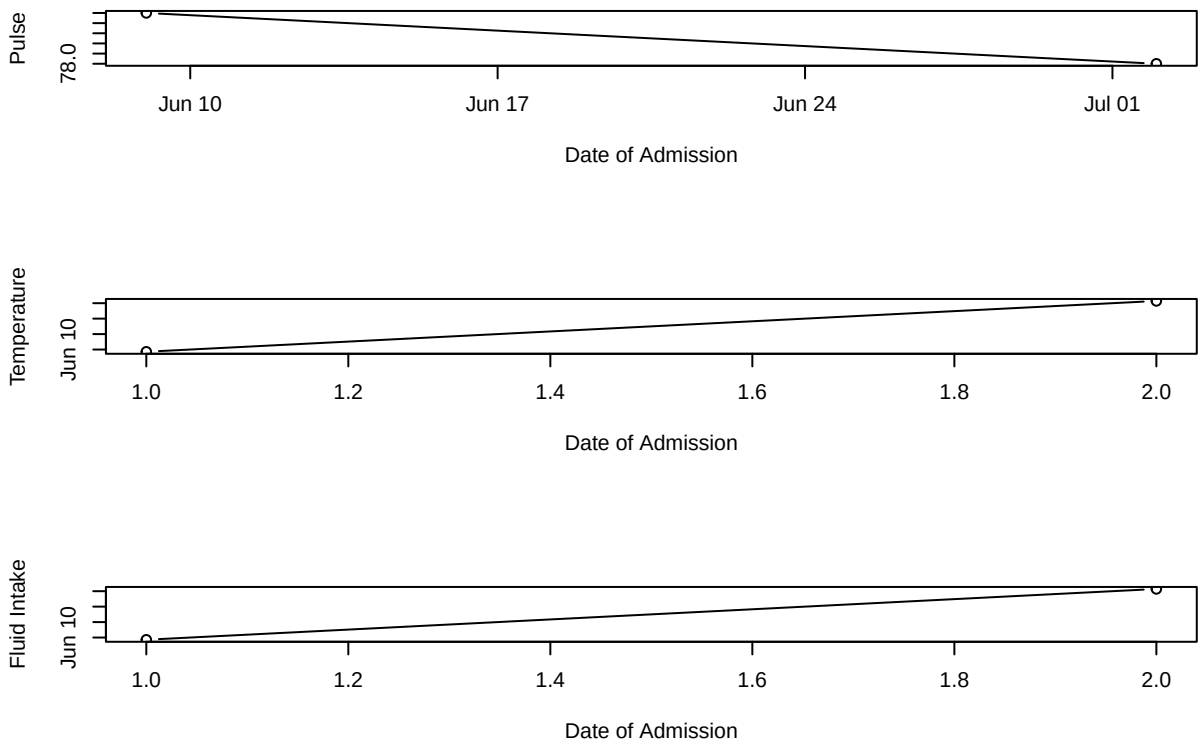
```
medicalRecord.mean(patient)
```

```
## $mean_pulse
## [1] 78.5
##
## $mean_temperature
## [1] 97.785
##
## $mean_fluid_intake
## [1] 0.4
```

```
medicalRecord.print(patient)
```

```
## Name: Yes
## Gender: male
## Date of Birth: 1977-05-03
##
## Test Results:
##   date_of_admission pulse temperature fluid_intake
## 1        2013-06-09    79       98.07         0.28
## 2        2013-07-02    78       97.50         0.52
```

```
medicalRecord.plot(patient)
```



3. Create a further class for a cohort (group) of patients, and write methods for `mean` and `print` which, when applied to a cohort, apply mean or print to each patient contained in the cohort. Hint: think of this as a "container" for patients. Reset the RNG seed to 8 and create a cohort of ten patients, then show the output for `mean` and `print`. (5 points)

```
set.seed(8)
n <- 10

medicalRecord <- function() {
  vowel <- grep("[aeiou]", letters)
  cons <- grep("[^aeiou]", letters)
  name <- paste(sample(LETTERS[cons], 1), sample(letters[vowel], 1), sample(letters[cons], 1), sep='')
  gender <- factor(sample(0:1, 1), levels=0:1, labels=c('female','male'))
  dob <- as.Date(sample(7500, 1), origin="1970-01-01")
  n <- sample(6, 1)
  doa <- as.Date(sample(1500, n), origin="2010-01-01")
  pulse <- round(rnorm(n, 80, 10))
  temp <- round(rnorm(n, 98.4, 0.3), 2)
  fluid <- round(runif(n), 2)
  list(name = name, gender = gender, dob = dob, doa = doa, pulse = pulse, temp = temp, fluid = fluid)
}

sample_cohort <- list()
```

```
  for (i in (1:n)) {
    patient <- medicalRecord()
    class(patient) <- "medicalRecord"
    sample_cohort[[i]] <- patient
  }

cohort <- function(x) {
  UseMethod("cohort")
}

cohort.mean <- function(x) {
  lapply(x,medicalRecord.mean)
}

cohort.print <- function(x) {
  lapply(x,medicalRecord.print)
}
```

```
cohort.mean(sample_cohort)
```

```
## [[1]]
## [[1]]$mean_pulse
## [1] 78.5
##
## [[1]]$mean_temperature
## [1] 97.785
##
## [[1]]$mean_fluid_intake
## [1] 0.4
##
##
## [[2]]
## [[2]]$mean_pulse
## [1] 86.33333
##
## [[2]]$mean_temperature
## [1] 98.39667
##
## [[2]]$mean_fluid_intake
## [1] 0.4133333
##
##
## [[3]]
## [[3]]$mean_pulse
## [1] 77
##
## [[3]]$mean_temperature
## [1] 98.6475
##
## [[3]]$mean_fluid_intake
## [1] 0.52
##
##
```

```
## [[4]]
## [[4]]$mean_pulse
## [1] 83.16667
##
## [[4]]$mean_temperature
## [1] 98.485
##
## [[4]]$mean_fluid_intake
## [1] 0.2966667
##
##
## [[5]]
## [[5]]$mean_pulse
## [1] 83.5
##
## [[5]]$mean_temperature
## [1] 98.45
##
## [[5]]$mean_fluid_intake
## [1] 0.4525
##
##
## [[6]]
## [[6]]$mean_pulse
## [1] 84.4
##
## [[6]]$mean_temperature
## [1] 98.484
##
## [[6]]$mean_fluid_intake
## [1] 0.522
##
##
## [[7]]
## [[7]]$mean_pulse
## [1] 76.5
##
## [[7]]$mean_temperature
## [1] 98.38
##
## [[7]]$mean_fluid_intake
## [1] 0.3975
##
##
## [[8]]
## [[8]]$mean_pulse
## [1] 75
##
## [[8]]$mean_temperature
## [1] 98.3675
##
## [[8]]$mean_fluid_intake
## [1] 0.5225
##
```

```
##
## [[9]]
## [[9]]$mean_pulse
## [1] 73
##
## [[9]]$mean_temperature
## [1] 98.36
##
## [[9]]$mean_fluid_intake
## [1] 0.15
##
##
## [[10]]
## [[10]]$mean_pulse
## [1] 77
##
## [[10]]$mean_temperature
## [1] 98.54
##
## [[10]]$mean_fluid_intake
## [1] 0.15
```

```
cohort.print(sample_cohort)
```

```
## Name: Yes
## Gender: male
## Date of Birth: 1977-05-03
##
## Test Results:
##   date_of_admission pulse temperature fluid_intake
## 1        2013-06-09    79       98.07         0.28
## 2        2013-07-02    78       97.50         0.52
## Name: Fal
## Gender: male
## Date of Birth: 1988-05-24
##
## Test Results:
##   date_of_admission pulse temperature fluid_intake
## 1        2010-11-16    76       98.23         0.18
## 2        2013-09-12    96       98.75         0.96
## 3        2013-03-24    87       98.21         0.10
## Name: Zog
## Gender: male
## Date of Birth: 1988-12-14
##
## Test Results:
##   date_of_admission pulse temperature fluid_intake
## 1        2013-03-25    69       98.49         0.81
## 2        2013-07-29    75       98.82         0.59
## 3        2013-10-27    80       98.74         0.28
## 4        2010-02-24    84       98.54         0.40
## Name: Yol
## Gender: male
## Date of Birth: 1986-03-11
```

```
##
## Test Results:
##   date_of_admission pulse temperature fluid_intake
## 1        2014-01-28    69       98.29         0.03
## 2        2013-03-24    78       98.44         0.13
## 3        2012-03-10    87       98.78         0.12
## 4        2010-02-22    84       98.87         0.39
## 5        2011-12-27    89       98.27         0.97
## 6        2012-11-26    92       98.26         0.14
## Name: Yak
## Gender: female
## Date of Birth: 1983-09-15
##
## Test Results:
##   date_of_admission pulse temperature fluid_intake
## 1        2012-08-30    90       98.58         0.26
## 2        2012-04-07    88       97.53         0.29
## 3        2011-07-19    75       98.58         0.60
## 4        2012-07-11    81       99.11         0.66
## Name: Gaf
## Gender: female
## Date of Birth: 1978-04-27
##
## Test Results:
##   date_of_admission pulse temperature fluid_intake
## 1        2012-04-24    89       98.32         0.42
## 2        2010-07-19    91       98.01         0.47
## 3        2012-08-06    77       98.96         0.74
## 4        2013-08-21    75       98.52         0.62
## 5        2011-05-03    90       98.61         0.36
## Name: Kuw
## Gender: female
## Date of Birth: 1980-11-07
##
## Test Results:
##   date_of_admission pulse temperature fluid_intake
## 1        2011-09-16    72       98.21         0.29
## 2        2010-10-29    81       98.17         0.93
## 3        2012-07-10    71       98.65         0.25
## 4        2010-10-03    82       98.49         0.12
## Name: Mav
## Gender: female
## Date of Birth: 1989-07-16
##
## Test Results:
##   date_of_admission pulse temperature fluid_intake
## 1        2012-03-02    63       99.07         0.01
## 2        2010-06-11    83       98.45         0.79
## 3        2010-02-08    66       97.95         0.79
## 4        2010-04-19    88       98.00         0.50
## Name: Fel
## Gender: male
## Date of Birth: 1985-08-16
##
```

```
## Test Results:
##   date_of_admission pulse temperature fluid_intake
## 1        2012-06-24    65       98.21         0.06
## 2        2010-09-26    81       98.51         0.24
## Name: Say
## Gender: female
## Date of Birth: 1974-09-22
##
## Test Results:
##   date_of_admission pulse temperature fluid_intake
## 1        2010-03-14    77       98.54         0.15

## [[1]]
##   date_of_admission pulse temperature fluid_intake
## 1        2013-06-09    79       98.07         0.28
## 2        2013-07-02    78       97.50         0.52
##
## [[2]]
##   date_of_admission pulse temperature fluid_intake
## 1        2010-11-16    76       98.23         0.18
## 2        2013-09-12    96       98.75         0.96
## 3        2013-03-24    87       98.21         0.10
##
## [[3]]
##   date_of_admission pulse temperature fluid_intake
## 1        2013-03-25    69       98.49         0.81
## 2        2013-07-29    75       98.82         0.59
## 3        2013-10-27    80       98.74         0.28
## 4        2010-02-24    84       98.54         0.40
##
## [[4]]
##   date_of_admission pulse temperature fluid_intake
## 1        2014-01-28    69       98.29         0.03
## 2        2013-03-24    78       98.44         0.13
## 3        2012-03-10    87       98.78         0.12
## 4        2010-02-22    84       98.87         0.39
## 5        2011-12-27    89       98.27         0.97
## 6        2012-11-26    92       98.26         0.14
##
## [[5]]
##   date_of_admission pulse temperature fluid_intake
## 1        2012-08-30    90       98.58         0.26
## 2        2012-04-07    88       97.53         0.29
## 3        2011-07-19    75       98.58         0.60
## 4        2012-07-11    81       99.11         0.66
##
## [[6]]
##   date_of_admission pulse temperature fluid_intake
## 1        2012-04-24    89       98.32         0.42
## 2        2010-07-19    91       98.01         0.47
## 3        2012-08-06    77       98.96         0.74
## 4        2013-08-21    75       98.52         0.62
## 5        2011-05-03    90       98.61         0.36
##
```

```
## [[7]]
##   date_of_admission pulse temperature fluid_intake
## 1        2011-09-16    72       98.21         0.29
## 2        2010-10-29    81       98.17         0.93
## 3        2012-07-10    71       98.65         0.25
## 4        2010-10-03    82       98.49         0.12
##
## [[8]]
##   date_of_admission pulse temperature fluid_intake
## 1        2012-03-02    63       99.07         0.01
## 2        2010-06-11    83       98.45         0.79
## 3        2010-02-08    66       97.95         0.79
## 4        2010-04-19    88       98.00         0.50
##
## [[9]]
##   date_of_admission pulse temperature fluid_intake
## 1        2012-06-24    65       98.21         0.06
## 2        2010-09-26    81       98.51         0.24
##
## [[10]]
##   date_of_admission pulse temperature fluid_intake
## 1        2010-03-14    77       98.54         0.15
```