Homework 2

Noah McIntire

Problem 1

a

```
commutes <- matrix(c(25,22,36,23,21,36,34,33,25,32), nrow=5, ncol=2, byrow=TRUE)
commutes
```

```
##      [,1] [,2]
## [1,]   25   22
## [2,]   36   23
## [3,]   21   36
## [4,]   34   33
## [5,]   25   32
```

b

```
rownames(commutes) <- c("Monday", "Tuesday", "Wednesday", "Thursday", "Friday")
colnames(commutes) <- c("Week1", "Week2")
commutes
```

```
##           Week1 Week2
## Monday       25    22
## Tuesday      36    23
## Wednesday    21    36
## Thursday     34    33
## Friday       25    32
```

c

On Monday, Tuesday, and Thursday she arrived faster when leaving at 8:30am. On Wednesday and Friday she arrived faster when leaving at 8:30am. ### d

```
avg_time<-apply(commutes, 1,mean)
avg_time
```

```
##    Monday   Tuesday Wednesday  Thursday    Friday
##      23.5      29.5      28.5      33.5      28.5
```

e

```
diff <- commutes -27
diff
```

```
##           Week1 Week2
## Monday       -2    -5
## Tuesday       9    -4
## Wednesday    -6     9
## Thursday      7     6
## Friday       -2     5
```

f

```
avg_diff<-apply(diff, 2,mean)
avg_diff
```

```
## Week1 Week2
##   1.2   2.2
```

g

```
max_delay<-apply(diff, 2,max)
max_delay
```

```
## Week1 Week2
##     9     9
```

h

```
Under_half<- commutes[commutes[,"Week2"] > 30,]
Under_half<-Under_half[,2]
Under_half<-names(Under_half)
Under_half
```

```
## [1] "Wednesday" "Thursday"  "Friday"
```

i

She arrived within her budgeted window on Monday, Wednesday, and Friday during the first week, and arrived within her window on Monday and Tuesday of the second week.

J

```
## She arrived fastest on Wednesday during the first week, which is the third row
fast<- rownames(commutes)
fast=fast[3]
fast
```

```
## [1] "Wednesday"
```

k

```
diff_sub <- diff[c(1,4),]
diff_sub
```

```
##          Week1 Week2
## Monday      -2    -5
## Thursday     7     6
```

Problem 2

a

```
library("car")
```

```
## Loading required package: carData
```

```
require("car")
weight.metric <-Davis[,c(2,4)]
head(weight.metric)
```

```
##   weight repwt
## 1     77    77
## 2     58    51
## 3     53    54
## 4     68    70
## 5     59    59
## 6     76    76
```

b

```
weight.imp <- weight.metric * 2.2
head(weight.imp)
```

```
##   weight repwt
## 1  169.4 169.4
## 2  127.6 112.2
## 3  116.6 118.8
## 4  149.6 154.0
## 5  129.8 129.8
## 6  167.2 167.2
```

c

```
height.metric <- Davis[,c(3,5)]
head(height.metric)
```

```
##   height repht
## 1    182   180
## 2    161   159
## 3    161   158
## 4    177   175
## 5    157   155
## 6    170   165
```

d

```
height.imp <- round(height.metric /2.54, 1)
head(height.imp)
```

```
##   height repht
## 1   71.7  70.9
## 2   63.4  62.6
## 3   63.4  62.2
## 4   69.7  68.9
## 5   61.8  61.0
## 6   66.9  65.0
```

e

```
Davis.imp <- data.frame(Davis$sex, weight.imp, height.imp )
colnames(Davis.imp)<- c("sex", "rec.weight",
"rep.weight", "rec.height", "rep.height")
head(Davis.imp)
```

```
##   sex rec.weight rep.weight rec.height rep.height
## 1   M      169.4      169.4       71.7       70.9
## 2   F      127.6      112.2       63.4       62.6
## 3   F      116.6      118.8       63.4       62.2
## 4   M      149.6      154.0       69.7       68.9
## 5   F      129.8      129.8       61.8       61.0
## 6   M      167.2      167.2       66.9       65.0
```

f

```
summary(Davis.imp)
```

```
##  sex       rec.weight      rep.weight      rec.height      rep.height
##  F:112   Min.   : 85.8   Min.   : 90.2   Min.   :22.40   Min.   :58.30
##  M: 88   1st Qu.:121.0   1st Qu.:121.0   1st Qu.:64.60   1st Qu.:63.20
##          Median :138.6   Median :138.6   Median :66.70   Median :66.10
##          Mean   :144.8   Mean   :144.4   Mean   :66.94   Mean   :66.34
##          3rd Qu.:162.8   3rd Qu.:161.7   3rd Qu.:69.80   3rd Qu.:68.90
##          Max.   :365.2   Max.   :272.8   Max.   :77.60   Max.   :78.70
##                          NA's   :17                      NA's   :17
```

There are 34 total NA values.

g

```
# From https://discuss.analyticsvidhya.com/t/how-to-count-the-missing-value-in-r/2949/5
rowSums(is.na(Davis.imp))
```

```
##   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20
##   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
##  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40
##   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
##  41  42  43  44  45  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
##   0   0   0   0   0   0   2   2   0   0   0   0   0   0   2   0   1   0   0   0
##  61  62  63  64  65  66  67  68  69  70  71  72  73  74  75  76  77  78  79  80
##   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   2   0   0   0   0
##  81  82  83  84  85  86  87  88  89  90  91  92  93  94  95  96  97  98  99 100
##   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   1
## 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
##   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
## 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140
##   0   0   0   0   2   1   1   0   0   0   0   0   0   0   0   0   0   2   0   0
## 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160
##   0   0   0   0   0   0   0   0   0   0   0   0   0   2   0   0   0   2   2   0
## 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
##   0   0   0   0   0   0   0   0   0   0   0   2   0   2   0   0   2   0   0   0
## 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200
##   0   2   2   0   0   0   0   0   0   0   0   0   0   0   0   0   0   2   0   0
```

There are 19 rows with values missing.

h

```
#From: https://stackoverflow.com/questions/7980622/subset-of-rows-containing-na-missing-
sex_subset <- Davis.imp[rowSums(is.na(Davis.imp)) > 0,]
sex_subset <- sex_subset$sex
sex_subset
```

```
##  [1] M F M F F F M F F F F F F F M F F M M
## Levels: F M
```

Problem 3

a

```
name<- c("Mercury","Venus","Earth","Mars","Jupiter","Saturn","Uranus","Neptune")
distance<-c(0.39,0.72,1,1.52,5.2,9.54,19.18,30.06)
type<-c("terrestrial","terrestrial","terrestrial","terrestrial","gas","gas","gas","gas")
diameter<-c(0.382,0.949,1,0.532,11.209,9.449,4.007,3.883)
rotation<-c(58.64,-243.02,1,1.03,0.41,0.43,-0.72,0.67)
rings<-c(F,F,F,F,T,T,T,T)
moons<-c(0,0,1,2,79,53,27,14)
solar_system<-data.frame(name, distance,
type, diameter, rotation, rings, moons)
solar_system
```

```
##      name distance        type diameter rotation rings moons
## 1 Mercury     0.39 terrestrial    0.382    58.64 FALSE     0
## 2   Venus     0.72 terrestrial    0.949  -243.02 FALSE     0
## 3   Earth     1.00 terrestrial    1.000     1.00 FALSE     1
## 4    Mars     1.52 terrestrial    0.532     1.03 FALSE     2
## 5 Jupiter     5.20         gas   11.209     0.41  TRUE    79
## 6  Saturn     9.54         gas    9.449     0.43  TRUE    53
## 7  Uranus    19.18         gas    4.007    -0.72  TRUE    27
## 8 Neptune    30.06         gas    3.883     0.67  TRUE    14
```

b

```
dia_l5<-solar_system[which(solar_system["diameter"] < 5),]
dia_l5
```

```
##      name distance        type diameter rotation rings moons
## 1 Mercury     0.39 terrestrial    0.382    58.64 FALSE     0
## 2   Venus     0.72 terrestrial    0.949  -243.02 FALSE     0
## 3   Earth     1.00 terrestrial    1.000     1.00 FALSE     1
## 4    Mars     1.52 terrestrial    0.532     1.03 FALSE     2
## 7  Uranus    19.18         gas    4.007    -0.72  TRUE    27
## 8 Neptune    30.06         gas    3.883     0.67  TRUE    14
```

c

```
neg_rot<-solar_system[which(solar_system["rotation"] < 0),]
neg_rot$distance
```

```
## [1]  0.72 19.18
```

d

```
dia_great<-solar_system[which(solar_system["diameter"] > 1),]
dia_great<-dia_great[,c(1,6,3)]
dia_great
```

```
##      name rings type
## 5 Jupiter  TRUE  gas
## 6  Saturn  TRUE  gas
## 7  Uranus  TRUE  gas
## 8 Neptune  TRUE  gas
```

e

```
moons_2<-solar_system[which(solar_system["moons"] > 1),]
moons_2<-moons_2[,c(6,2)]
moons_2
```

```
##   rings distance
## 4 FALSE     1.52
## 5  TRUE     5.20
## 6  TRUE     9.54
## 7  TRUE    19.18
## 8  TRUE    30.06
```

References

1. https://www.statmethods.net/input/missingdata.html
2. https://discuss.analyticsvidhya.com/t/how-to-count-the-missing-value-in-r/2949/5
3. https://stackoverflow.com/questions/7980622/subset-of-rows-containing-na-missing-values-in-a-chosen-column-of-a-data-frame