STAT 351 Homework 6

Charles Hwang

4/13/2020

Problem 1

```
rm(list=ls())
data <- read.csv(file="/Users/newuser/Desktop/Notes/Undergraduate/STAT 351 - Nonparametric Statistical I
library(lattice)
library(mice)
library(VIM)
names(data)              # Problem 1a
```

```
##  [1] "Country"                          "Region"
##  [3] "DataYear"                         "ClassGrade"
##  [5] "Gender"                           "Ageyears"
##  [7] "Handed"                           "Height_cm"
##  [9] "Footlength_cm"                    "Armspan_cm"
## [11] "Languages_spoken"                 "Travel_to_School"
## [13] "Travel_time_to_School"            "Reaction_time"
## [15] "Score_in_memory_game"             "Favourite_physical_activity"
## [17] "Importance_reducing_pollution"    "Importance_recycling_rubbish"
## [19] "Importance_conserving_water"      "Importance_saving_energy"
## [21] "Importance_owning_computer"       "Importance_Internet_access"
## [23] "Left_Footlength_cm"               "Longer_foot"
## [25] "Index_Fingerlength_mm"            "Ring_Fingerlength_mm"
## [27] "Longer_Finger_Lefthand"           "Birth_month"
## [29] "Favorite_Season"                  "Allergies"
## [31] "Vegetarian"                       "Favorite_Food"
## [33] "Beverage"                         "Favorite_School_Subject"
## [35] "Sleep_Hours_Schoolnight"          "Sleep_Hours_Non_Schoolnight"
## [37] "Home_Occupants"                   "Home_Internet_Access"
## [39] "Communication_With_Friends"       "Text_Messages_Sent_Yesterday"
## [41] "Text_Messages_Received_Yesterday" "Hanging_Out_With_Friends_Hours"
## [43] "Talking_On_Phone_Hours"           "Doing_Homework_Hours"
## [45] "Doing_Things_With_Family_Hours"   "Outdoor_Activities_Hours"
## [47] "Video_Games_Hours"                "Social_Websites_Hours"
## [49] "Texting_Messaging_Hours"          "Computer_Use_Hours"
## [51] "Watching_TV_Hours"                "Paid_Work_Hours"
## [53] "Work_At_Home_Hours"               "Schoolwork_Pressure"
## [55] "Planned_Education_Level"          "Favorite_Music"
## [57] "Superpower"                       "Preferred_Status"
## [59] "Role_Model_Type"                  "Charity_Donation"
```

```
data$Height_cm <- as.numeric(data$Height_cm)
data$Armspan_cm <- as.numeric(data$Armspan_cm)
table(data$ClassGrade)
```

```
##
##   9  10  11  12
## 117  99  12 272
```

```
table(data$Height_cm)
```

```
##
##   1.8     5   5.2    10    12  12.7    26    54    57    58 59.25    60  61.5
##     1     3     1     1     1     1     1     1     1     1     1     2     1
##    63  63.5    64    68    69    70    71    72    73    75  75.6  78.4    80
##     3     1     5     2     1     2     1     1     1     1     1     1     2
##    98   120   122   130   140   143   144   147 147.3   149 149.8 149.9   150
##     1     1     1     1     1     1     1     2     1     2     1     1     3
##   151   152 152.4   153   154 154.9   155   156   157 157.3 157.4 157.5   158
##     2     2     7     3     1     2     9     4     7     1     4     3     7
## 158.4 158.5   159   160 160.2   161 161.3   162 162.5 162.6 162.7   163 163.5
##     1     1     1    25     1     3     1    14     7     1     1    19     1
##   164   165 165.1   166 166.2   167 167.6   168   169   170 170.1 170.2 170.6
##     6    30     6     5     1    12     4    11     5    29     1     3     1
## 170.7   171   172 172.5 172.7   173   174   175 175.3   176 176.5 176.7   177
##     1     8    12     1     4     9     4    10     3    11     1     1    10
## 177.8   178   179   180 180.3   181   182 182.5 182.8   183 183.8   184   185
##     2     9     7     8     2     4     7     1     3     9     1     2     6
## 185.4 185.6   186   187 187.5   188   190 190.5   192   193   196   197   200
##     1     1     4     5     1     2     1     4     1     2     2     1     1
##   330   999 99999
##     1     1     1
```

```
# data[order(data$Vegetarian,data$Favorite_Food,decreasing=c(TRUE,FALSE))[1:sum(data$Vegetarian=="Yes")
levels(data$Favorite_Food)
```

```
## NULL
```

```
cat("The data set appears to describe ",nrow(data)," different high school students in Illinois. There a
```

```
## The data set appears to describe 500 different high school students in Illinois. There are many diffe
##
## There are several interesting characteristics in the data set, including that the median recorded hei
```

```
sum(is.na(data$Handed))  # Problem 1b
```

```
## [1] 0
```

```
sum(is.na(data$Favorite_Season)) # It is okay to remove the blanks because there are only two variables
```

```
## [1] 0
```

```
chisq.test(data$Handed,data$Favorite_Season,correct=FALSE) # Chi-squared test with original full data s
```

```
##
##  Pearson's Chi-squared test
##
## data:  data$Handed and data$Favorite_Season
## X-squared = 96.46, df = 12, p-value = 2.741e-15
```

```
chisq.test(data$Handed[data$Handed!=""&data$Favorite_Season!=""],data$Favorite_Season[data$Handed!=""&da
```

```
##
##  Pearson's Chi-squared test
```

```
##
## data:  data$Handed[data$Handed != "" & data$Favorite_Season != ""] and data$Favorite_Season[data$Hand
##     ""]
## X-squared = 3.7881, df = 6, p-value = 0.7053
```

```
cat("We fail to reject H0 at the a = .05 level. There is insufficient evidence (p = ",chisq.test(data$Ha
```

```
## We fail to reject H0 at the a = .05 level. There is insufficient evidence (p = 0.705328) that handedn
```

```
summary(data)            # Problem 1c
```

```
##    Country             Region            DataYear      ClassGrade
##  Length:500         Length:500         Min.   :2012   Min.   : 9.00
##  Class :character   Class :character   1st Qu.:2015   1st Qu.:10.00
##  Mode  :character   Mode  :character   Median :2017   Median :12.00
##                                        Mean   :2017   Mean   :10.88
##                                        3rd Qu.:2018   3rd Qu.:12.00
##                                        Max.   :2019   Max.   :12.00
##
##     Gender             Ageyears        Handed            Height_cm
##  Length:500         Min.   :12.00   Length:500         Min.   :    1.8
##  Class :character   1st Qu.:15.00   Class :character   1st Qu.:  160.0
##  Mode  :character   Median :17.00   Mode  :character   Median :  166.6
##                     Mean   :16.36                      Mean   :  377.0
##                     3rd Qu.:17.00                      3rd Qu.:  175.2
##                     Max.   :99.00                      Max.   :99999.0
##                     NA's   :4                          NA's   :38
##  Footlength_cm        Armspan_cm     Languages_spoken Travel_to_School
##  Length:500         Min.   :  1.88   Min.   : 0.00    Length:500
##  Class :character   1st Qu.:154.00   1st Qu.: 1.00    Class :character
##  Mode  :character   Median :164.00   Median : 2.00    Mode  :character
##                     Mean   :153.39   Mean   : 1.69
##                     3rd Qu.:175.00   3rd Qu.: 2.00
##                     Max.   :431.80   Max.   :13.00
##                     NA's   :55       NA's   :22
##  Travel_time_to_School Reaction_time      Score_in_memory_game
##  Length:500            Length:500         Length:500
##  Class :character      Class :character   Class :character
##  Mode  :character      Mode  :character   Mode  :character
##
##
##
##
##  Favourite_physical_activity Importance_reducing_pollution
##  Length:500                  Length:500
##  Class :character            Class :character
##  Mode  :character            Mode  :character
##
##
##
##
##  Importance_recycling_rubbish Importance_conserving_water
##  Min.   :   0.0               Min.   :   0.0
##  1st Qu.: 439.0               1st Qu.: 400.5
##  Median : 600.0               Median : 652.0
```

```
##  Mean   : 631.3               Mean   : 628.9
##  3rd Qu.: 850.0               3rd Qu.: 925.0
##  Max.   :9050.0               Max.   :5000.0
##  NA's   :47                   NA's   :49
##  Importance_saving_energy Importance_owning_computer Importance_Internet_access
##  Length:500               Length:500                 Length:500
##  Class :character         Class :character           Class :character
##  Mode  :character         Mode  :character           Mode  :character
##
##
##
##
##  Left_Footlength_cm Longer_foot        Index_Fingerlength_mm
##  Length:500         Length:500         Length:500
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##  Ring_Fingerlength_mm Longer_Finger_Lefthand Birth_month
##  Length:500           Length:500             Length:500
##  Class :character     Class :character       Class :character
##  Mode  :character     Mode  :character       Mode  :character
##
##
##
##
##  Favorite_Season     Allergies          Vegetarian         Favorite_Food
##  Length:500         Length:500         Length:500         Length:500
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##    Beverage         Favorite_School_Subject Sleep_Hours_Schoolnight
##  Length:500         Length:500              Length:500
##  Class :character   Class :character        Class :character
##  Mode  :character   Mode  :character        Mode  :character
##
##
##
##
##  Sleep_Hours_Non_Schoolnight Home_Occupants     Home_Internet_Access
##  Length:500                  Length:500         Length:500
##  Class :character            Class :character   Class :character
##  Mode  :character            Mode  :character   Mode  :character
##
##
##
##
##  Communication_With_Friends Text_Messages_Sent_Yesterday
##  Length:500                 Length:500
```

```
##  Class :character           Class :character
##  Mode  :character           Mode  :character
##
##
##
##
##  Text_Messages_Received_Yesterday Hanging_Out_With_Friends_Hours
##  Length:500                       Length:500
##  Class :character                 Class :character
##  Mode  :character                 Mode  :character
##
##
##
##
##  Talking_On_Phone_Hours Doing_Homework_Hours Doing_Things_With_Family_Hours
##  Length:500             Length:500           Length:500
##  Class :character       Class :character     Class :character
##  Mode  :character       Mode  :character     Mode  :character
##
##
##
##
##  Outdoor_Activities_Hours Video_Games_Hours  Social_Websites_Hours
##  Length:500               Length:500         Length:500
##  Class :character         Class :character   Class :character
##  Mode  :character         Mode  :character   Mode  :character
##
##
##
##
##  Texting_Messaging_Hours Computer_Use_Hours Watching_TV_Hours
##  Length:500              Length:500         Length:500
##  Class :character        Class :character   Class :character
##  Mode  :character        Mode  :character   Mode  :character
##
##
##
##
##  Paid_Work_Hours    Work_At_Home_Hours Schoolwork_Pressure
##  Length:500         Length:500         Length:500
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##  Planned_Education_Level Favorite_Music      Superpower
##  Length:500              Length:500         Length:500
##  Class :character        Class :character   Class :character
##  Mode  :character        Mode  :character   Mode  :character
##
##
##
##
```

```
##  Preferred_Status   Role_Model_Type    Charity_Donation
##  Length:500         Length:500         Length:500
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
##
##
##
##
```

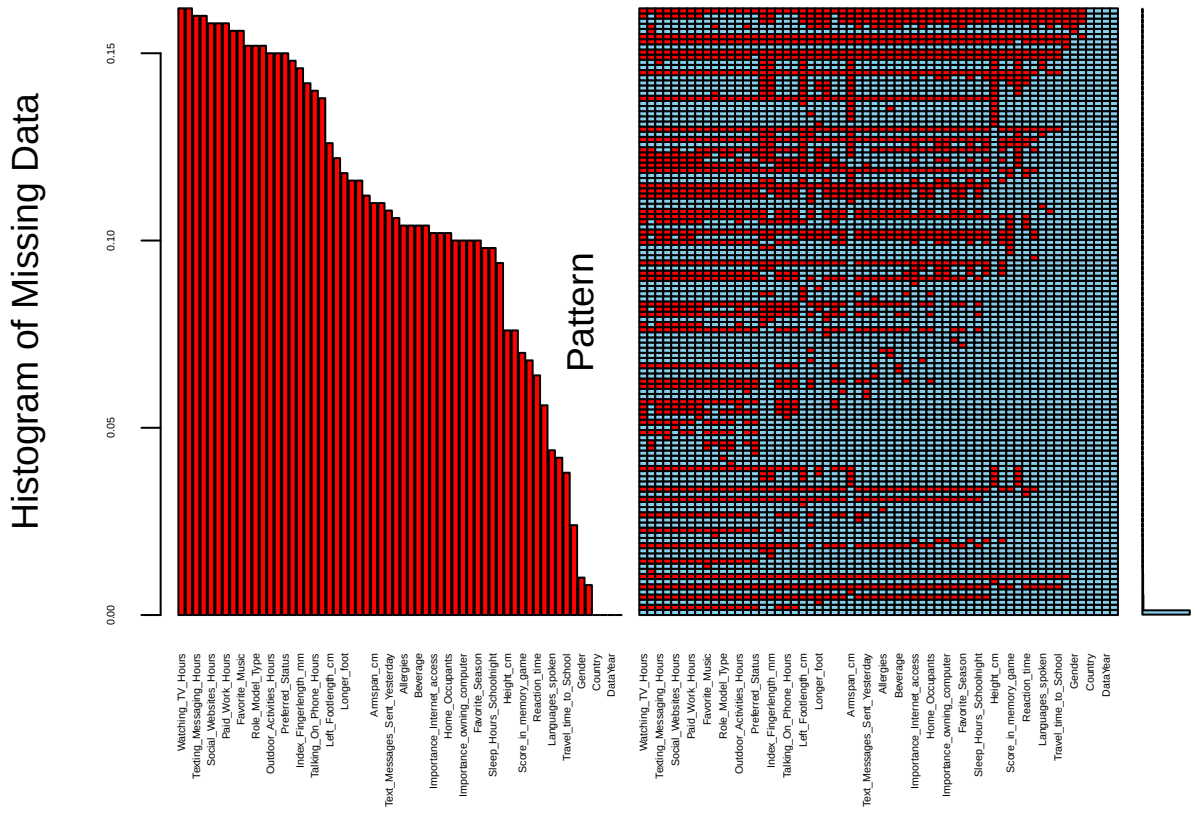
```
data[data==""] <- NA # Changing blank ("") values to "NA"
sum(is.na(data[,c("Height_cm","Armspan_cm")])) # There are no missing values in either the predictor or
```

```
## [1] 93
```

```
sum(is.na(data)) # However, there are 2,998 missing values in the rest of the data. Let's visualize tha
```

```
## [1] 3091
```

```
aggr(data,sortVars=TRUE,labels=names(data),ylab=c("Histogram of Missing Data","Pattern"),cex.axis=.35,g
```



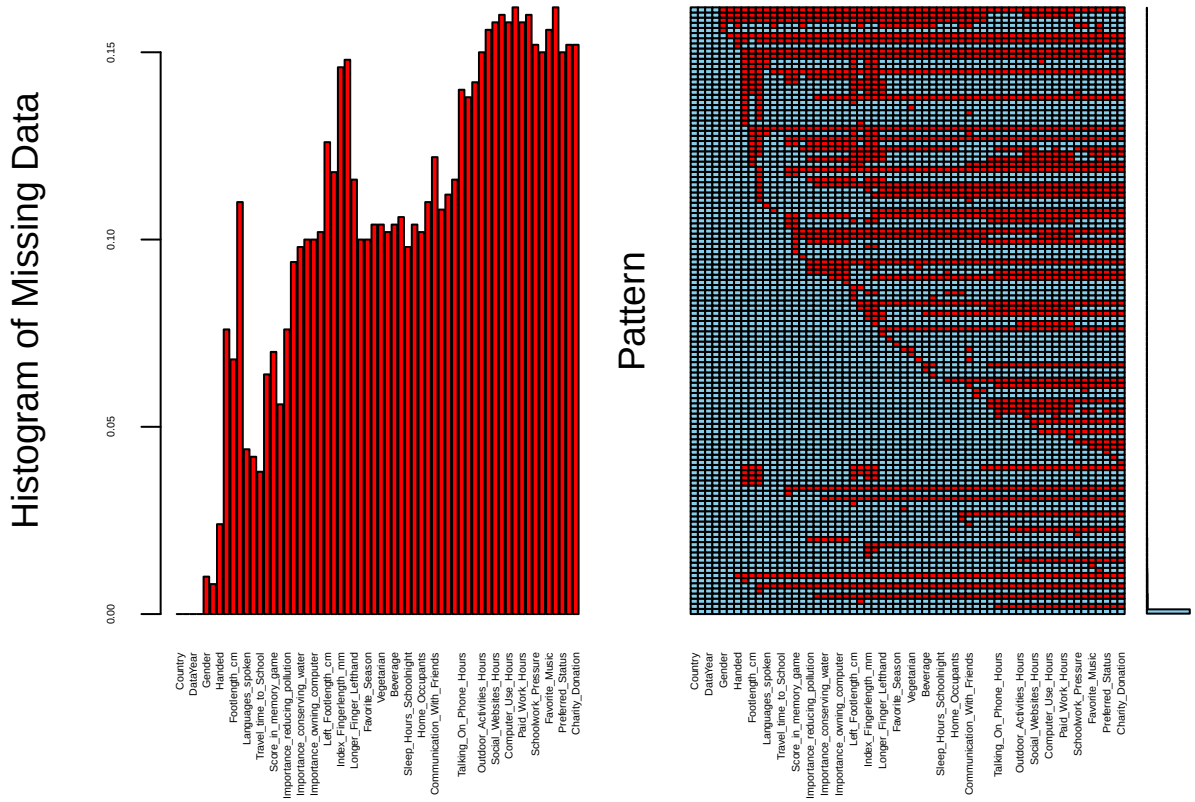

```
##
##  Variables sorted by number of missings:
##                        Variable Count
##               Watching_TV_Hours 0.162
##                      Superpower 0.162
##          Texting_Messaging_Hours 0.160
##               Work_At_Home_Hours 0.160
##            Social_Websites_Hours 0.158
##               Computer_Use_Hours 0.158
##                  Paid_Work_Hours 0.158
```

```
##                 Video_Games_Hours 0.156
##                    Favorite_Music 0.156
##               Schoolwork_Pressure 0.152
##                   Role_Model_Type 0.152
##                  Charity_Donation 0.152
##          Outdoor_Activities_Hours 0.150
##           Planned_Education_Level 0.150
##                  Preferred_Status 0.150
##              Ring_Fingerlength_mm 0.148
##             Index_Fingerlength_mm 0.146
##    Doing_Things_With_Family_Hours 0.142
##            Talking_On_Phone_Hours 0.140
##              Doing_Homework_Hours 0.138
##                Left_Footlength_cm 0.126
##        Communication_With_Friends 0.122
##                       Longer_foot 0.118
##            Longer_Finger_Lefthand 0.116
##    Hanging_Out_With_Friends_Hours 0.116
##  Text_Messages_Received_Yesterday 0.112
##                        Armspan_cm 0.110
##              Home_Internet_Access 0.110
##      Text_Messages_Sent_Yesterday 0.108
##           Favorite_School_Subject 0.106
##                         Allergies 0.104
##                        Vegetarian 0.104
##                          Beverage 0.104
##       Sleep_Hours_Non_Schoolnight 0.104
##        Importance_Internet_access 0.102
##                     Favorite_Food 0.102
##                    Home_Occupants 0.102
##          Importance_saving_energy 0.100
##        Importance_owning_computer 0.100
##                       Birth_month 0.100
##                    Favorite_Season 0.100
##       Importance_conserving_water 0.098
##           Sleep_Hours_Schoolnight 0.098
##      Importance_recycling_rubbish 0.094
##                        Height_cm 0.076
##     Importance_reducing_pollution 0.076
##               Score_in_memory_game 0.070
##                      Footlength_cm 0.068
##                     Reaction_time 0.064
##       Favourite_physical_activity 0.056
##                  Languages_spoken 0.044
##                  Travel_to_School 0.042
##             Travel_time_to_School 0.038
##                            Handed 0.024
##                            Gender 0.010
##                          Ageyears 0.008
##                           Country 0.000
##                            Region 0.000
##                          DataYear 0.000
##                       ClassGrade 0.000
```

```
cat("The table tells us that '",aggr(data,plot=FALSE,bars=FALSE)$missing$Variable[order(aggr(data,plot=
```

```
## The table tells us that 'Watching_TV_Hours' and 'Superpower' are the variables with the highest prop
```

```
aggr(data,labels=names(data),ylab=c("Histogram of Missing Data","Pattern"),cex.axis=.35,gap=2.5)
```




```
cat("The same histogram and plot with the original data set (sorted by variable position) shows that th
```

```
## The same histogram and plot with the original data set (sorted by variable position) shows that there
```

```
sum(is.na(data[,c(1:4)])) # Check
```

```
## [1] 0
```

```
cart <- mice(data[,c("Height_cm","Armspan_cm","Gender","Ageyears","Handed","Footlength_cm","Languages_s
```

```
##
##  iter imp variable
##    1   1  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    1   2  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    1   3  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    1   4  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    1   5  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    2   1  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    2   2  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    2   3  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    2   4  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    2   5  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    3   1  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    3   2  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    3   3  Height_cm  Armspan_cm  Ageyears  Languages_spoken
```

```
##   3   4  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##   3   5  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##   4   1  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##   4   2  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##   4   3  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##   4   4  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##   4   5  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##   5   1  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##   5   2  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##   5   3  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##   5   4  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##   5   5  Height_cm  Armspan_cm  Ageyears  Languages_spoken
```

```
summary(cart)
```

```
## Class: mids
## Number of multiple imputations:  5
## Imputation methods:
##                   Height_cm                  Armspan_cm
##                      "cart"                      "cart"
##                      Gender                     Ageyears
##                          ""                        "cart"
##                      Handed                Footlength_cm
##                          ""                          ""
##           Languages_spoken       Travel_time_to_School
##                      "cart"                          ""
## Favourite_physical_activity                 Longer_foot
##                          ""                          ""
##      Longer_Finger_Lefthand                 Birth_month
##                          ""                          ""
##             Favorite_Season                    Allergies
##                          ""                          ""
##                    Beverage     Favorite_School_Subject
##                          ""                          ""
##     Sleep_Hours_Schoolnight              Home_Occupants
##                          ""                          ""
##        Home_Internet_Access         Schoolwork_Pressure
##                          ""                          ""
##     Planned_Education_Level              Favorite_Music
##                          ""                          ""
##                  Superpower             Role_Model_Type
##                          ""                          ""
##           Charity_Donation
##                          ""
## PredictorMatrix:
##               Height_cm Armspan_cm Gender Ageyears Handed Footlength_cm
## Height_cm             0          1      0        1      0             0
## Armspan_cm            1          0      0        1      0             0
## Gender                0          0      0        0      0             0
## Ageyears              1          1      0        0      0             0
## Handed                0          0      0        0      0             0
## Footlength_cm         0          0      0        0      0             0
##               Languages_spoken Travel_time_to_School
## Height_cm                    1                     0
## Armspan_cm                   1                     0
```

```
## Gender                       0                    0
## Ageyears                     1                    0
## Handed                       0                    0
## Footlength_cm                0                    0
##               Favourite_physical_activity Longer_foot Longer_Finger_Lefthand
## Height_cm                               0           0                      0
## Armspan_cm                              0           0                      0
## Gender                                  0           0                      0
## Ageyears                                0           0                      0
## Handed                                  0           0                      0
## Footlength_cm                           0           0                      0
##               Birth_month Favorite_Season Allergies Beverage
## Height_cm               0               0         0        0
## Armspan_cm              0               0         0        0
## Gender                  0               0         0        0
## Ageyears                0               0         0        0
## Handed                  0               0         0        0
## Footlength_cm           0               0         0        0
##               Favorite_School_Subject Sleep_Hours_Schoolnight Home_Occupants
## Height_cm                           0                       0              0
## Armspan_cm                          0                       0              0
## Gender                              0                       0              0
## Ageyears                            0                       0              0
## Handed                              0                       0              0
## Footlength_cm                       0                       0              0
##               Home_Internet_Access Schoolwork_Pressure Planned_Education_Level
## Height_cm                        0                   0                       0
## Armspan_cm                       0                   0                       0
## Gender                           0                   0                       0
## Ageyears                         0                   0                       0
## Handed                           0                   0                       0
## Footlength_cm                    0                   0                       0
##               Favorite_Music Superpower Role_Model_Type Charity_Donation
## Height_cm                 0          0               0                0
## Armspan_cm                0          0               0                0
## Gender                    0          0               0                0
## Ageyears                  0          0               0                0
## Handed                    0          0               0                0
## Footlength_cm             0          0               0                0
## Number of logged events:  21
##   it im dep     meth                       out
## 1  0  0     constant                    Gender
## 2  0  0     constant                    Handed
## 3  0  0     constant              Footlength_cm
## 4  0  0     constant       Travel_time_to_School
## 5  0  0     constant Favourite_physical_activity
## 6  0  0     constant                 Longer_foot
```

```
sapply(complete(cart),function(x) sum(is.na(x))) # Checking that missing values were imputed
```

```
##                   Height_cm                  Armspan_cm
##                           0                           0
##                      Gender                     Ageyears
##                           5                           0
##                      Handed                Footlength_cm
```

```
##                           12                           34
##            Languages_spoken       Travel_time_to_School
##                            0                           19
## Favourite_physical_activity                Longer_foot
##                           28                           59
##      Longer_Finger_Lefthand                Birth_month
##                           58                           50
##             Favorite_Season                   Allergies
##                           50                           52
##                    Beverage     Favorite_School_Subject
##                           52                           53
##     Sleep_Hours_Schoolnight              Home_Occupants
##                           49                           51
##        Home_Internet_Access         Schoolwork_Pressure
##                           55                           76
##     Planned_Education_Level              Favorite_Music
##                           75                           78
##                  Superpower             Role_Model_Type
##                           81                           76
##            Charity_Donation
##                           76
```

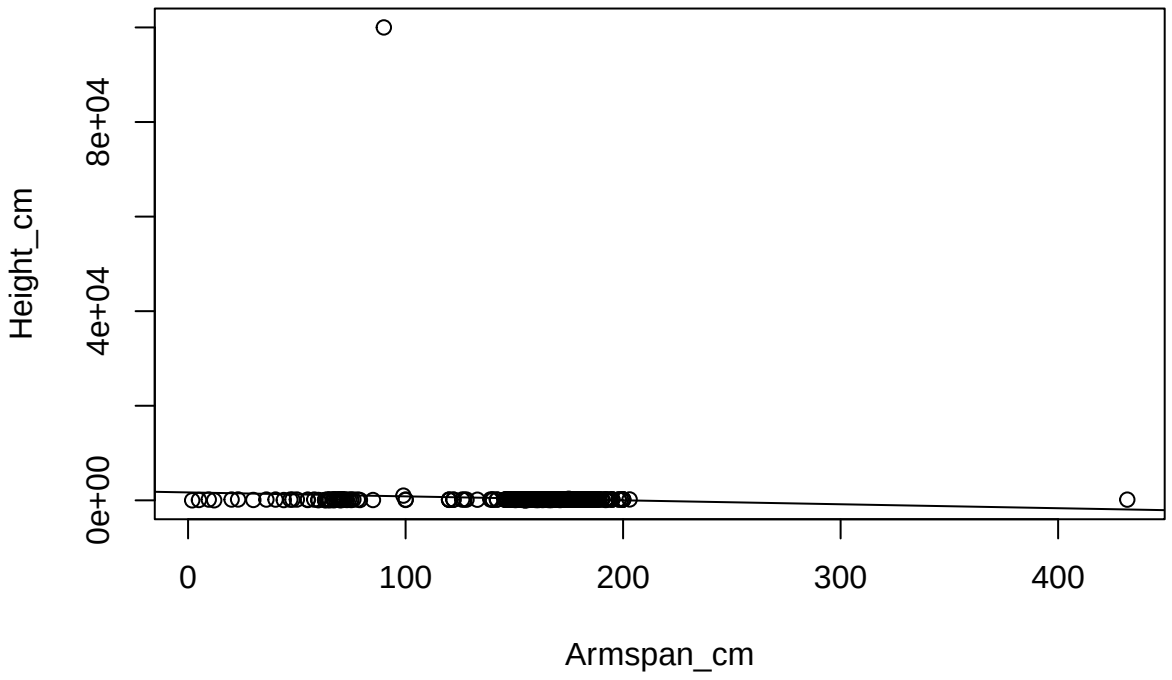
```
summary(pool(with(cart,lm(Height_cm~Armspan_cm,data=data))))
```

```
##          term    estimate  std.error statistic       df    p.value
## 1 (Intercept) 1675.900461 883.647665  1.896571 434.9697 0.05854687
## 2  Armspan_cm   -8.351213   5.542058 -1.506879 434.9697 0.13256748
```

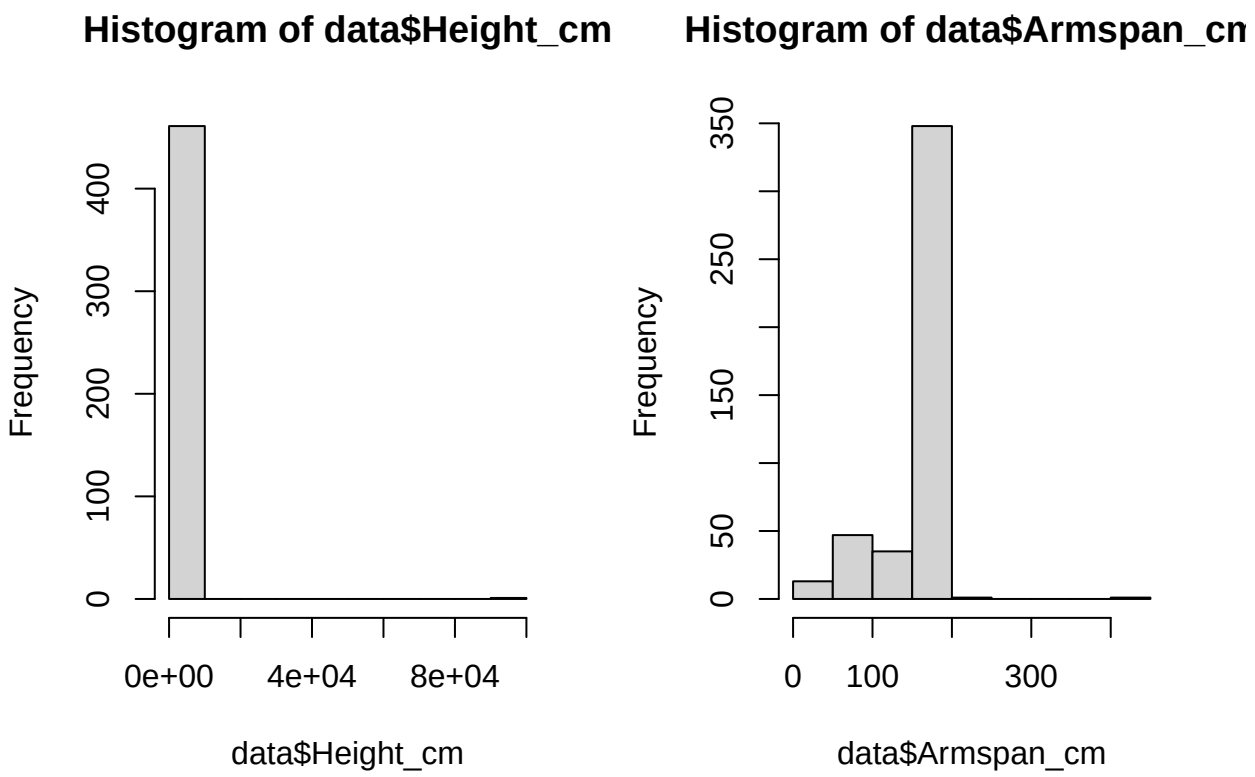
```
cat("The final estimates for the slope and intercept are ",summary(pool(with(cart,lm(Height_cm~Armspan_
```

```
## The final estimates for the slope and intercept are 1675.9 and -8.351213, respectively. The standard
```

```
plot(data[c("Armspan_cm","Height_cm")]) # This is how the line would appear if we were to plot height v
abline(coef=summary(pool(with(cart,lm(Height_cm~Armspan_cm,data=data))))$estimate)
```



```r
par(mfrow=c(1,2)) # However, these estimates really should be taken with a grain of salt. As I mentione
hist(data$Height_cm)
abline(v=median(data$Height_cm))
hist(data$Armspan_cm)
abline(v=median(data$Armspan_cm))
```




```r
data <- droplevels(data) # Problem 1d
rf <- mice(data[,c("Height_cm","Armspan_cm","Gender","Ageyears","Handed","Footlength_cm","Languages_spo
```

```
##
##  iter imp variable
##    1   1  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    1   2  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    1   3  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    1   4  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    1   5  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    2   1  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    2   2  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    2   3  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    2   4  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    2   5  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    3   1  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    3   2  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    3   3  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    3   4  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    3   5  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    4   1  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    4   2  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##    4   3  Height_cm  Armspan_cm  Ageyears  Languages_spoken
```

```
##   4   4  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##   4   5  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##   5   1  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##   5   2  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##   5   3  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##   5   4  Height_cm  Armspan_cm  Ageyears  Languages_spoken
##   5   5  Height_cm  Armspan_cm  Ageyears  Languages_spoken
```

```
summary(rf)
```

```
## Class: mids
## Number of multiple imputations:  5
## Imputation methods:
##                   Height_cm                  Armspan_cm
##                        "rf"                        "rf"
##                      Gender                     Ageyears
##                          ""                        "rf"
##                     Handed                Footlength_cm
##                          ""                          ""
##           Languages_spoken       Travel_time_to_School
##                       "rf"                          ""
## Favourite_physical_activity                 Longer_foot
##                          ""                          ""
##     Longer_Finger_Lefthand                 Birth_month
##                          ""                          ""
##             Favorite_Season                   Allergies
##                          ""                          ""
##                    Beverage     Favorite_School_Subject
##                          ""                          ""
##    Sleep_Hours_Schoolnight              Home_Occupants
##                          ""                          ""
##       Home_Internet_Access         Schoolwork_Pressure
##                          ""                          ""
##    Planned_Education_Level               Favorite_Music
##                          ""                          ""
##                 Superpower             Role_Model_Type
##                          ""                          ""
##           Charity_Donation
##                          ""
## PredictorMatrix:
##               Height_cm Armspan_cm Gender Ageyears Handed Footlength_cm
## Height_cm             0          1      0        1      0             0
## Armspan_cm            1          0      0        1      0             0
## Gender                0          0      0        0      0             0
## Ageyears              1          1      0        0      0             0
## Handed                0          0      0        0      0             0
## Footlength_cm         0          0      0        0      0             0
##               Languages_spoken Travel_time_to_School
## Height_cm                    1                    0
## Armspan_cm                   1                    0
## Gender                       0                    0
## Ageyears                     1                    0
## Handed                       0                    0
## Footlength_cm                0                    0
##               Favourite_physical_activity Longer_foot Longer_Finger_Lefthand
```

```
## Height_cm                             0           0                     0
## Armspan_cm                            0           0                     0
## Gender                                0           0                     0
## Ageyears                              0           0                     0
## Handed                                0           0                     0
## Footlength_cm                         0           0                     0
##              Birth_month Favorite_Season Allergies Beverage
## Height_cm              0               0         0        0
## Armspan_cm             0               0         0        0
## Gender                 0               0         0        0
## Ageyears               0               0         0        0
## Handed                 0               0         0        0
## Footlength_cm          0               0         0        0
##              Favorite_School_Subject Sleep_Hours_Schoolnight Home_Occupants
## Height_cm                          0                       0              0
## Armspan_cm                         0                       0              0
## Gender                             0                       0              0
## Ageyears                           0                       0              0
## Handed                             0                       0              0
## Footlength_cm                      0                       0              0
##              Home_Internet_Access Schoolwork_Pressure Planned_Education_Level
## Height_cm                       0                   0                       0
## Armspan_cm                      0                   0                       0
## Gender                          0                   0                       0
## Ageyears                        0                   0                       0
## Handed                          0                   0                       0
## Footlength_cm                   0                   0                       0
##              Favorite_Music Superpower Role_Model_Type Charity_Donation
## Height_cm                 0          0               0                0
## Armspan_cm                0          0               0                0
## Gender                    0          0               0                0
## Ageyears                  0          0               0                0
## Handed                    0          0               0                0
## Footlength_cm             0          0               0                0
## Number of logged events:  21
##   it im dep     meth                       out
## 1  0  0     constant                    Gender
## 2  0  0     constant                    Handed
## 3  0  0     constant             Footlength_cm
## 4  0  0     constant       Travel_time_to_School
## 5  0  0     constant Favourite_physical_activity
## 6  0  0     constant                Longer_foot
```

```
sapply(complete(rf),function(x) sum(is.na(x))) # Checking that missing values were imputed
```

```
##                   Height_cm                  Armspan_cm
##                           0                           0
##                      Gender                    Ageyears
##                           5                           0
##                     Handed                Footlength_cm
##                          12                          34
##            Languages_spoken       Travel_time_to_School
##                           0                          19
## Favourite_physical_activity                 Longer_foot
##                          28                          59
```

```
##      Longer_Finger_Lefthand                 Birth_month
##                          58                          50
##             Favorite_Season                   Allergies
##                          50                          52
##                    Beverage     Favorite_School_Subject
##                          52                          53
##     Sleep_Hours_Schoolnight              Home_Occupants
##                          49                          51
##        Home_Internet_Access         Schoolwork_Pressure
##                          55                          76
##     Planned_Education_Level              Favorite_Music
##                          75                          78
##                  Superpower             Role_Model_Type
##                          81                          76
##            Charity_Donation
##                          76
```

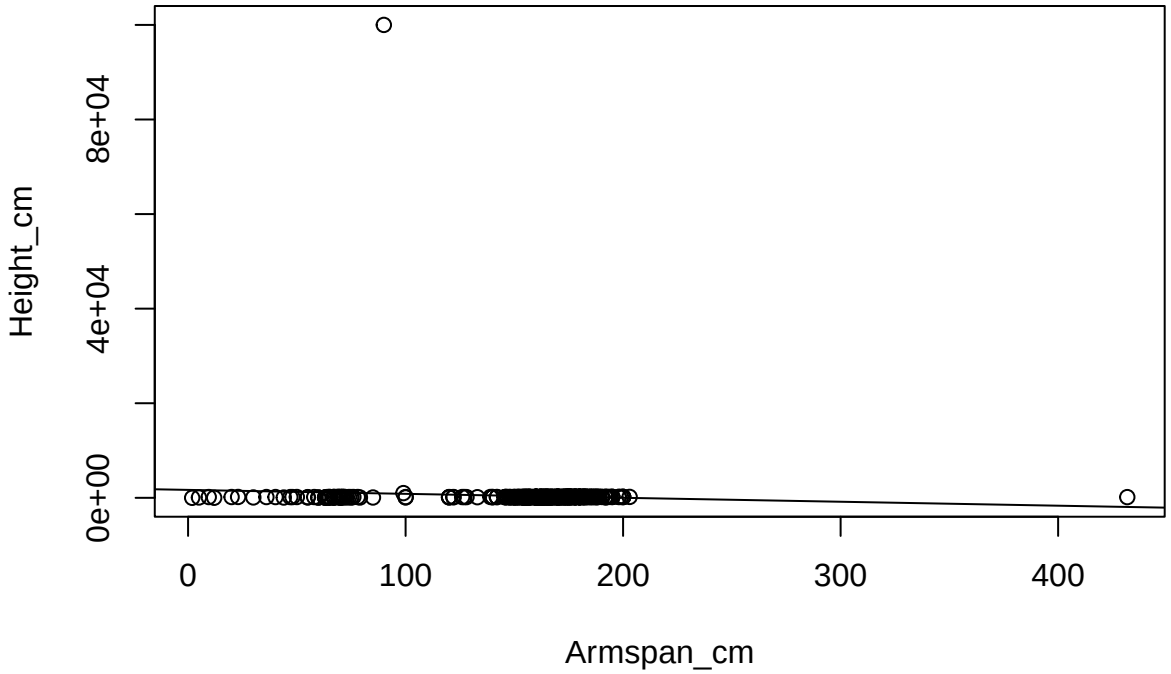
```
summary(pool(with(rf,lm(Height_cm~Armspan_cm,data=data))))
```

```
##          term    estimate  std.error statistic       df    p.value
## 1 (Intercept) 1675.900461 883.647665  1.896571 434.9697 0.05854687
## 2  Armspan_cm   -8.351213   5.542058 -1.506879 434.9697 0.13256748
```

```
cat("The final estimates for the slope and intercept are ",summary(pool(with(rf,lm(Height_cm~Armspan_cm
```

```
## The final estimates for the slope and intercept are 1675.9 and -8.351213, respectively. The standard
```

```
par(mfrow=c(1,1)) # This is how the line would appear if we were to plot height versus armspan on a gra
plot(data[c("Armspan_cm","Height_cm")])
abline(coef=summary(pool(with(rf,lm(Height_cm~Armspan_cm,data=data))))$estimate)
```




```
library(utils)          # Problem 1e
browseURL("https://github.com/LoyolaRamblers/Undergraduate-Coursework/tree/master/STAT%20351%20Homework
```