HW#5 87 F21 part 2

Lucía Carrera

10/29/2021

Question 2 – Measurement data

part a – reading in the data and 'cleaning' it: 1 pt

```
knitr::opts_chunk$set(echo = TRUE)

#  Load the tidyverse (or ggplot2 and dplyr) below:
library(tidyverse)

#  Download the MeasureF21.csv data; put it in your Stat/CS 087 project folder; rea
d it like this:
m <- read.csv('MeasureF21.csv', stringsAsFactors = TRUE,
              na.strings = '')

tibble(m)
```

```
## # A tibble: 55 x 10
##    ï..Response_id ID    Height Armspan  Head Rhand Lhand RFoot LFoot Hand
##             <int> <fct>  <dbl>   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <fct>
##  1        1199767 727     61.8    62.5  22    7.5   7.75   8.5  8.5  Right
##  2        1199768 CAL3    64.8    65.5  22.5  8.5   8.25   9.5 10.5  Right
##  3        1199769 6101    68.2    68    23.5  7.5   7.5   10.2 10.2  Right
##  4        1199770 J0HN    70.2    73    23.5  8.5   8.5   10.5 10.5  Right
##  5        1199771 6434    70      68    23    8.25  8.5   10.5 10.5  Right
##  6        1199772 3717    69      68.8  22    8     8.25  10   10    Right
##  7        1199773 AB10    67.2    68.8  23.2  8     8     10    9    Right
##  8        1199774 YPIA    70.2    70.5  23    8.75  8.5   10.5 10.8  Right
##  9        1199775 8888    66.5    67.5  23    7.5   6.5    9.5  9.75 Right
## 10        1199776 6542    69.5    65    22.5  8.5   8.25  10   10    Right
## # ... with 45 more rows
```

```
names(m)[1] <- "Response_id" # response id appeared in a strage way

#  Data cleaning done in class must be included here:
mnew <- m %>% filter(Response_id != 1199794,
                     Head > 20) %>%
  mutate(Response_id = as.factor(Response_id),
         RFoot = ifelse(RFoot > 20, RFoot/2.54, RFoot),
         LFoot = ifelse(LFoot > 20, LFoot/2.54, LFoot))
```

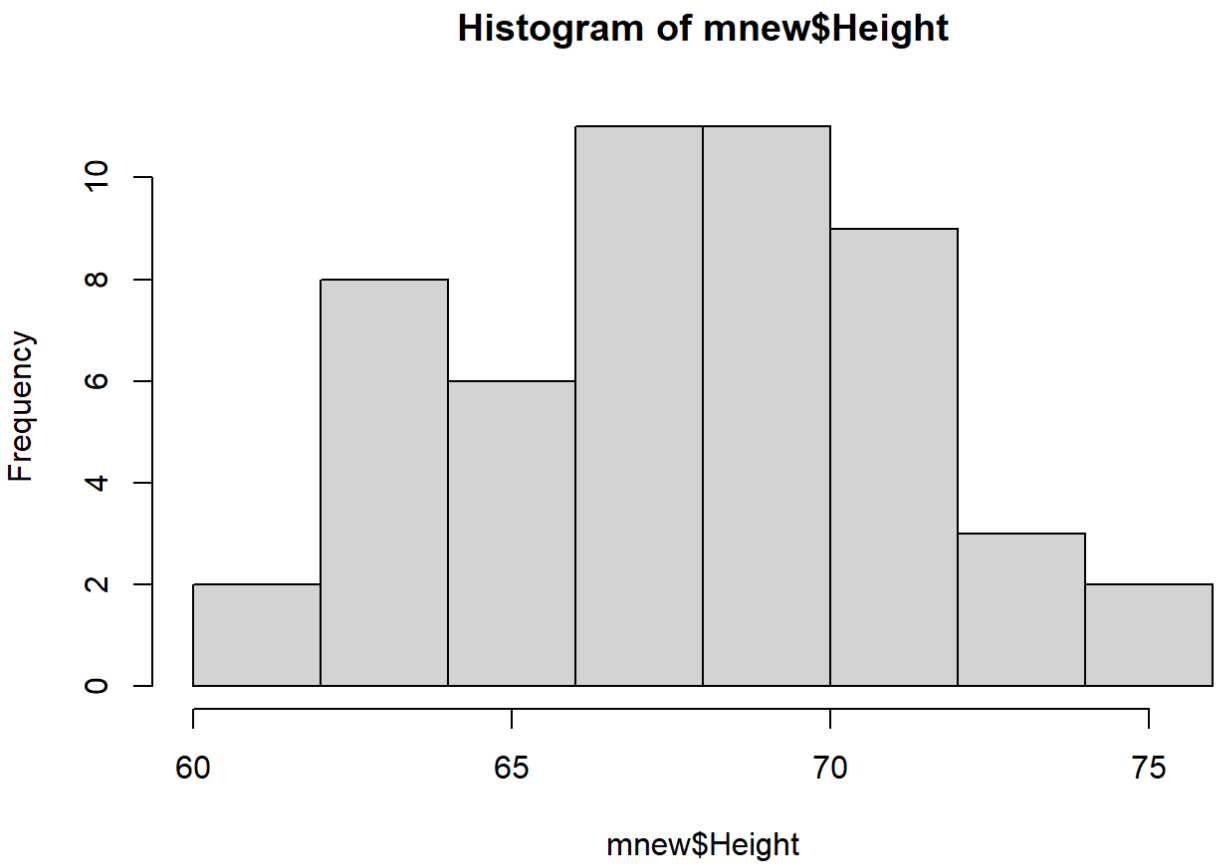
```
#  Print the data file (it's not too long), and make sure it is 'cleaned'.
mnew
```

```
##    Response_id   ID Height Armspan  Head Rhand Lhand RFoot LFoot          Hand
## 1      1199767  727  61.75   62.50 22.00  7.50  7.75  8.50  8.50         Right
## 2      1199768 CAL3  64.75   65.50 22.50  8.50  8.25  9.50 10.50         Right
## 3      1199769 6101  68.25   68.00 23.50  7.50  7.50 10.25 10.25         Right
## 4      1199770 J0HN  70.25   73.00 23.50  8.50  8.50 10.50 10.50         Right
## 5      1199771 6434  70.00   68.00 23.00  8.25  8.50 10.50 10.50         Right
## 6      1199772 3717  69.00   68.75 22.00  8.00  8.25 10.00 10.00         Right
## 7      1199773 AB10  67.25   68.75 23.25  8.00  8.00 10.00  9.00         Right
## 8      1199774 YPIA  70.25   70.50 23.00  8.75  8.50 10.50 10.75         Right
## 9      1199775 8888  66.50   67.50 23.00  7.50  6.50  9.50  9.75         Right
## 10     1199776 6542  69.50   65.00 22.50  8.50  8.25 10.00 10.00         Right
## 11     1199777 7402  72.25   74.25 22.75  8.50  8.25 10.75 11.00         Right
## 12     1199778 LENA  62.50   62.00 23.00  8.00  8.00  9.00  9.00         Right
## 13     1199779 7535  63.50   61.50 22.00  7.50  7.75  9.00  9.50          Left
## 14     1199780 cOoL  72.00   72.00 23.00  8.00  8.25  9.50 10.00         Right
## 15     1199781 5713  64.00   64.00 21.50  7.00  7.00  9.50  9.50         Right
## 16     1199782 9163  72.75   71.25 23.00  9.25  9.25 11.00 10.75         Right
## 17     1199783 rs24  62.50   60.50 22.50  8.00  8.25  9.40  9.60         Right
## 18     1199784 1019  67.25   66.00 23.25  7.50  7.75  9.75 10.00         Right
## 19     1199785 5734  69.50   69.50 22.75  8.75  8.75 10.50 10.50         Right
## 20     1199787 1369  67.50   64.50 22.00  8.50  8.75 10.50 10.50         Right
## 21     1199788  504  74.50   78.25 23.50  9.00  8.75 11.25 11.50         Right
## 22     1199789 1112  70.00   70.00 23.00  8.25  8.25 10.50 10.25         Right
## 23     1199790  312  69.50   70.50 22.00  8.75  9.00 10.75 11.00         Right
## 24     1199791 LA99  64.00   64.00 22.00  8.00  8.00 10.00  9.50          Left
## 25     1199792 8944  66.75   66.50 23.50  7.75  8.00 10.00 10.00         Right
## 26     1199793 CDB3  69.50   69.50 22.25  7.75  7.75  9.25  9.25         Right
## 27     1199795 7B59  67.00   63.50 22.50  7.75  8.00  9.00  9.25         Right
## 28     1199821 roww  70.00   72.00 22.50  8.25  8.25 10.50 10.75         Right
## 29     1199822 5678  67.00   63.00 23.00  7.00  7.00  9.00  9.25         Right
## 30     1199823 EA36  75.50   76.50 22.00  8.75  8.50 10.50 10.50          Left
## 31     1199824 9153  63.00   64.50 23.00  8.25  8.25  9.00  9.00         Right
## 32     1199825 5555  66.00      NA 21.00    NA    NA    NA  9.00          <NA>
## 33     1199826 1007  68.50   68.00 22.00  7.75  8.00  9.75 10.00          Left
## 34     1199827 1220  62.00   61.00 21.00  7.00  7.50  9.00  9.00         Right
## 35     1199828 3ena  63.00   62.75 22.00  7.75  7.75  8.75  9.00         Right
## 36     1199829 439b  72.00   72.00 23.50  7.65  7.80 11.60 11.70         Right
## 37     1199830 4765  65.00   65.50 22.10  7.50  7.60  9.50  9.50         Right
## 38     1199831 307N  73.50   76.00 22.75  8.75  9.25 11.00 11.25         Right
## 39     1199832 3741  64.50   64.00 21.75  7.50  7.50  9.00  9.00         Right
## 40     1199833  706  64.50   64.00 22.00  7.25  7.25  9.00  9.00         Right
## 41     1199834 BC69  71.00   71.00 22.00  8.00  8.00 10.50 10.50 Ambidexterous
## 42     1199835 3141  68.00   67.25 23.25  7.75  8.00 10.00 10.00         Right
## 43     1199836  310  63.00   66.00 22.75  9.00  8.75 10.50 10.75         Right
## 44     1199837 CHCH  67.00   71.25 23.75  9.00  9.00 10.50 11.00         Right
## 45     1199838 BEL3  71.75   77.50 22.50  9.00  9.25 11.25 10.50         Right
## 46     1199839 6429  68.00   71.00 22.00  8.25  8.25 10.50 10.50         Right
## 47     1199840 B3AN  68.50   69.50 23.25  9.75  9.75 10.50 10.50         Right
## 48     1199841 ETN7  70.50   71.50 22.00  8.00  8.00  9.25  9.75         Right
## 49     1199842 FM70  66.50   66.00 24.25  8.00  8.00  9.75  9.25         Right
## 50     1199843 LUCI  66.00   66.00 23.00  7.00  7.00    NA    NA         Right
## 51     1199844 rawr  71.25   75.00 23.00  8.00  8.00 10.50 10.25         Right
## 52     1199845 M1C7  71.00   74.00 23.50  8.25  8.00 10.75 10.50          Left
```

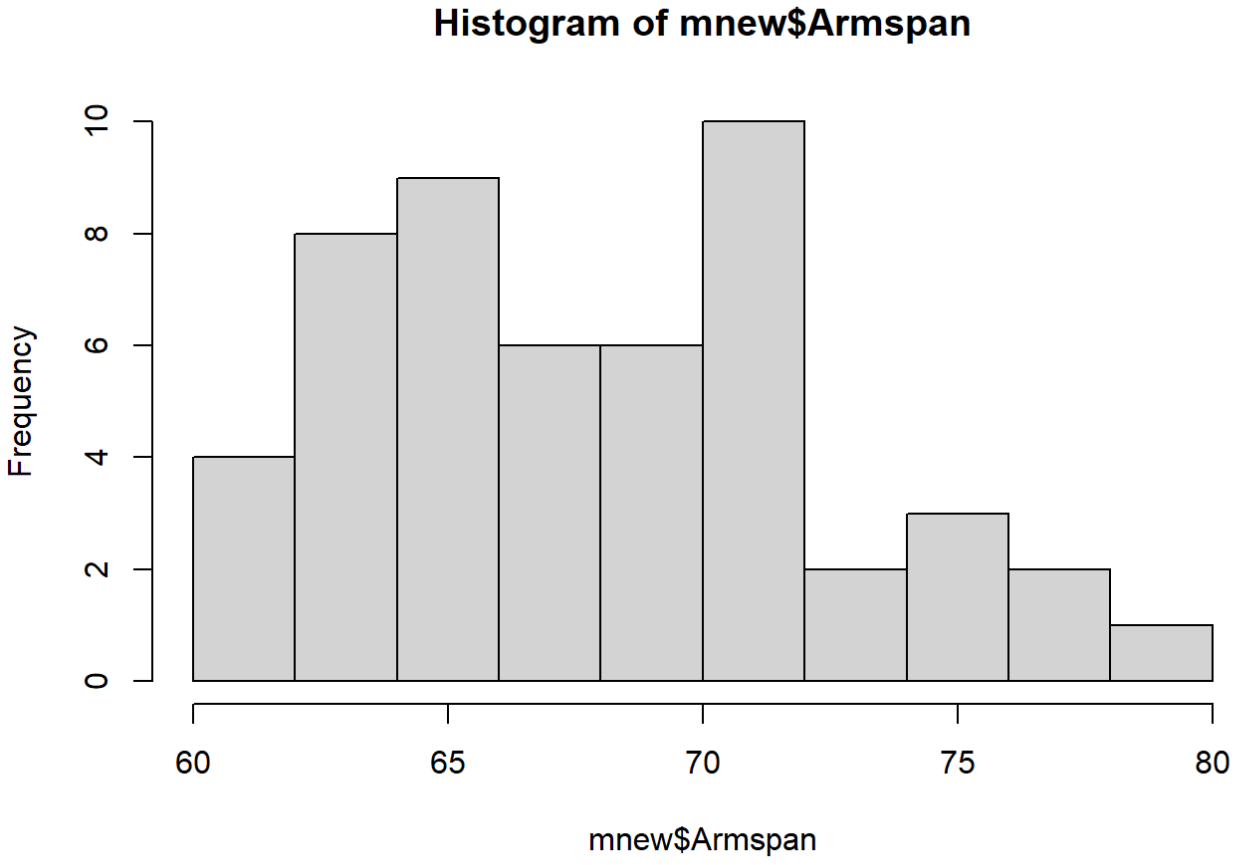
```
#  Do a summary of the data
summary(mnew)
```

```
##   Response_id       ID          Height          Armspan           Head
##  1199767: 1   1007   : 1   Min.   :61.75   Min.   :60.50   Min.   :21.00
##  1199768: 1   1019   : 1   1st Qu.:64.94   1st Qu.:64.50   1st Qu.:22.00
##  1199769: 1   1112   : 1   Median :68.00   Median :68.00   Median :22.75
##  1199770: 1   1220   : 1   Mean   :67.88   Mean   :68.25   Mean   :22.63
##  1199771: 1   1369   : 1   3rd Qu.:70.25   3rd Qu.:71.25   3rd Qu.:23.00
##  1199772: 1   307N   : 1   Max.   :75.50   Max.   :78.25   Max.   :24.25
##  (Other):46   (Other):46                   NA's   :1
##      Rhand           Lhand           RFoot            LFoot
##  Min.   :7.000   Min.   :6.500   Min.   : 8.500   Min.   : 8.500
##  1st Qu.:7.700   1st Qu.:7.750   1st Qu.: 9.425   1st Qu.: 9.375
##  Median :8.000   Median :8.000   Median :10.000   Median :10.000
##  Mean   :8.091   Mean   :8.125   Mean   : 9.990   Mean   :10.025
##  3rd Qu.:8.500   3rd Qu.:8.500   3rd Qu.:10.500   3rd Qu.:10.500
##  Max.   :9.750   Max.   :9.750   Max.   :11.600   Max.   :11.700
##  NA's   :1       NA's   :1       NA's   :2        NA's   :1
##             Hand
##  Ambidexterous: 1
##  Left         : 5
##  Right        :45
##  NA's         : 1
##
##
##
```

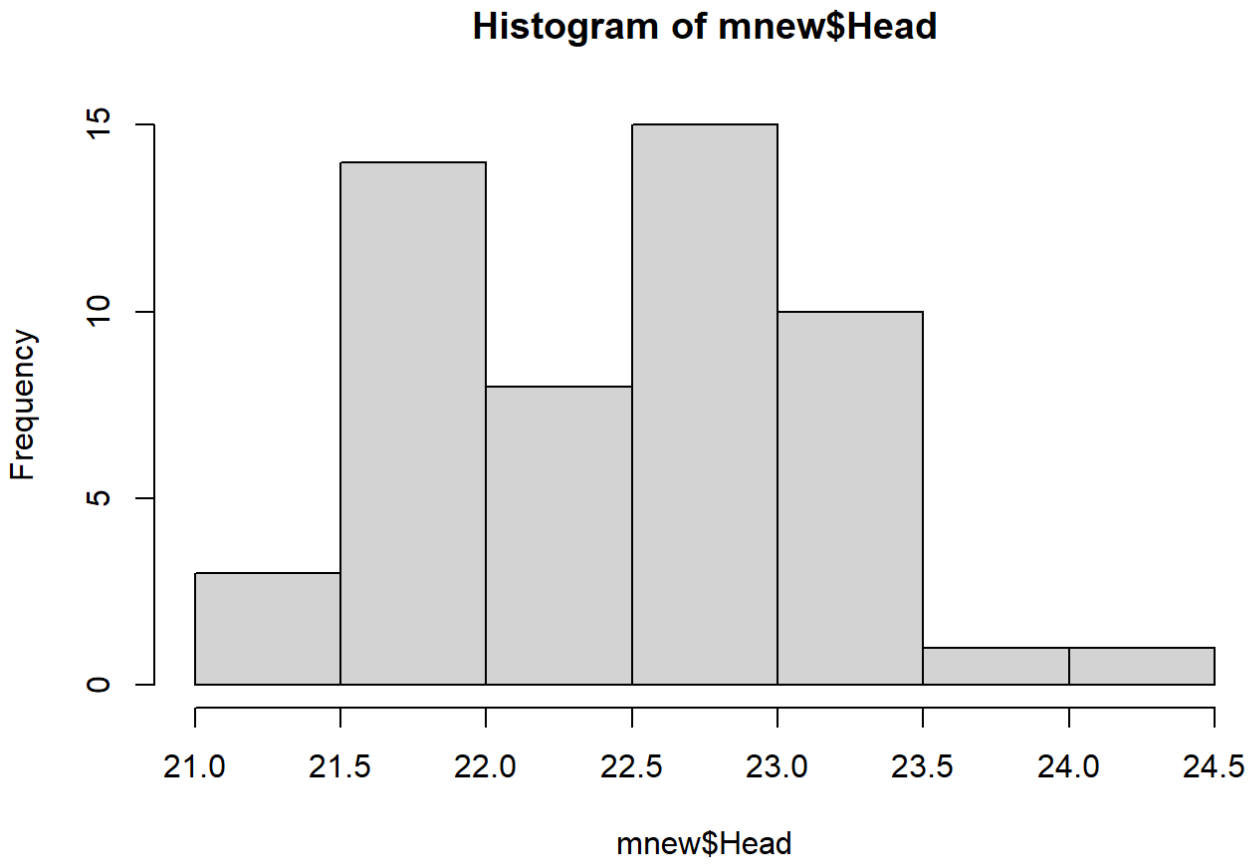
```
hist(mnew$Height)
```



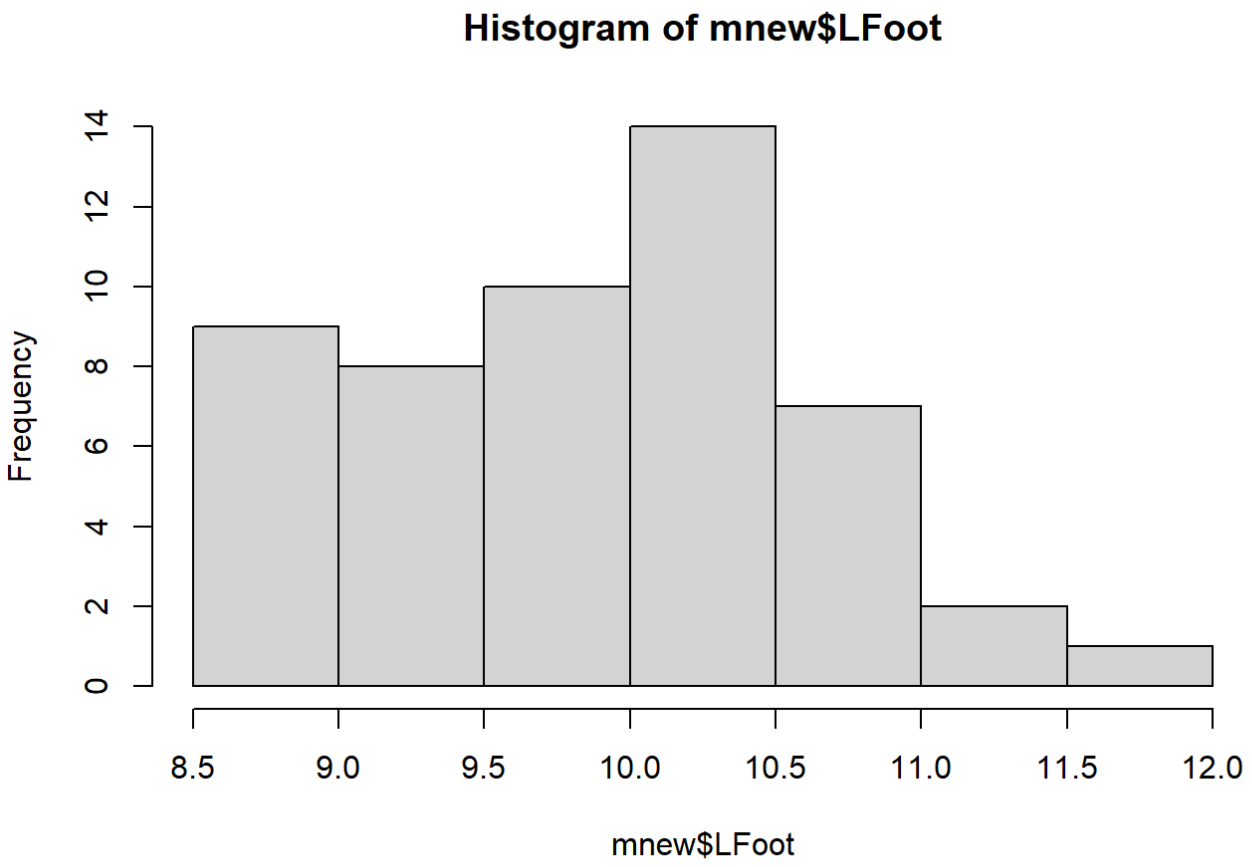

```
hist(mnew$Armspan)
```



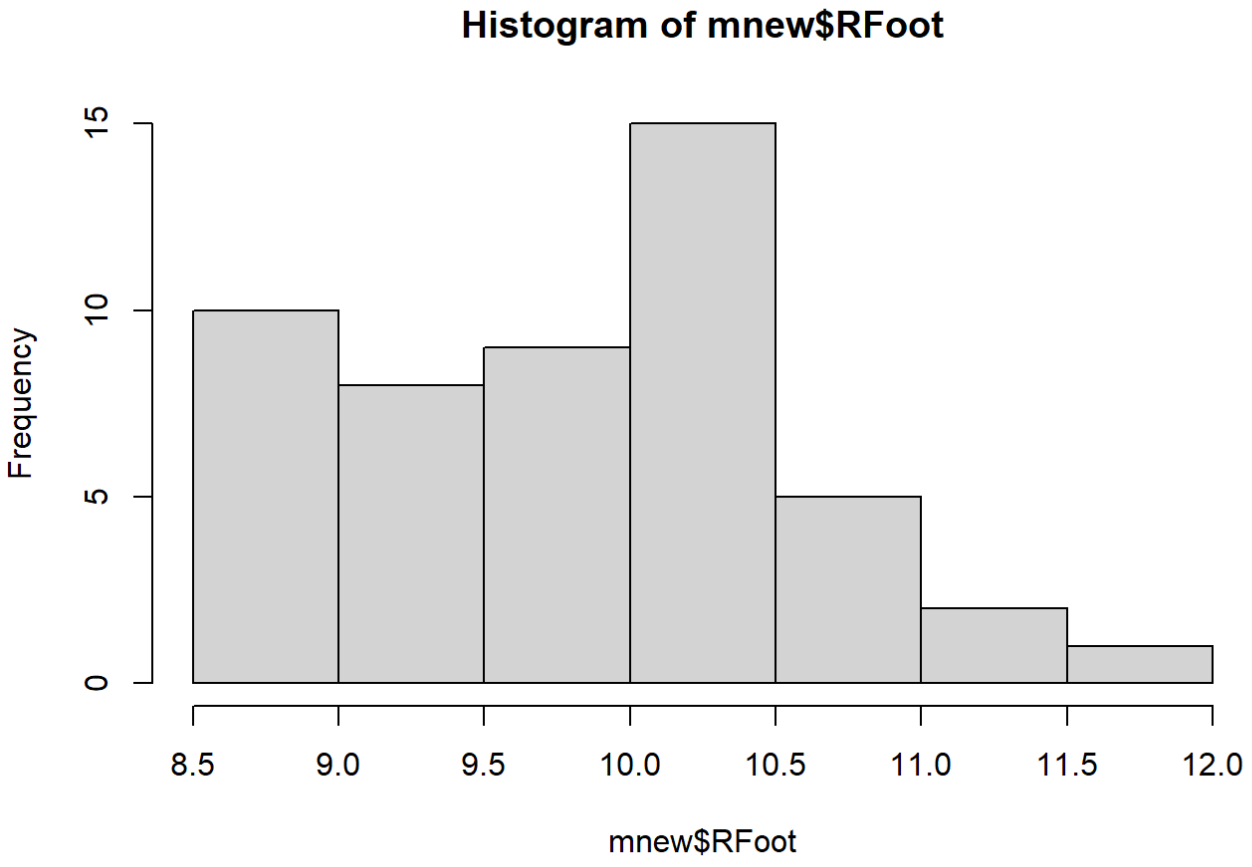
```
hist(mnew$Head)
```



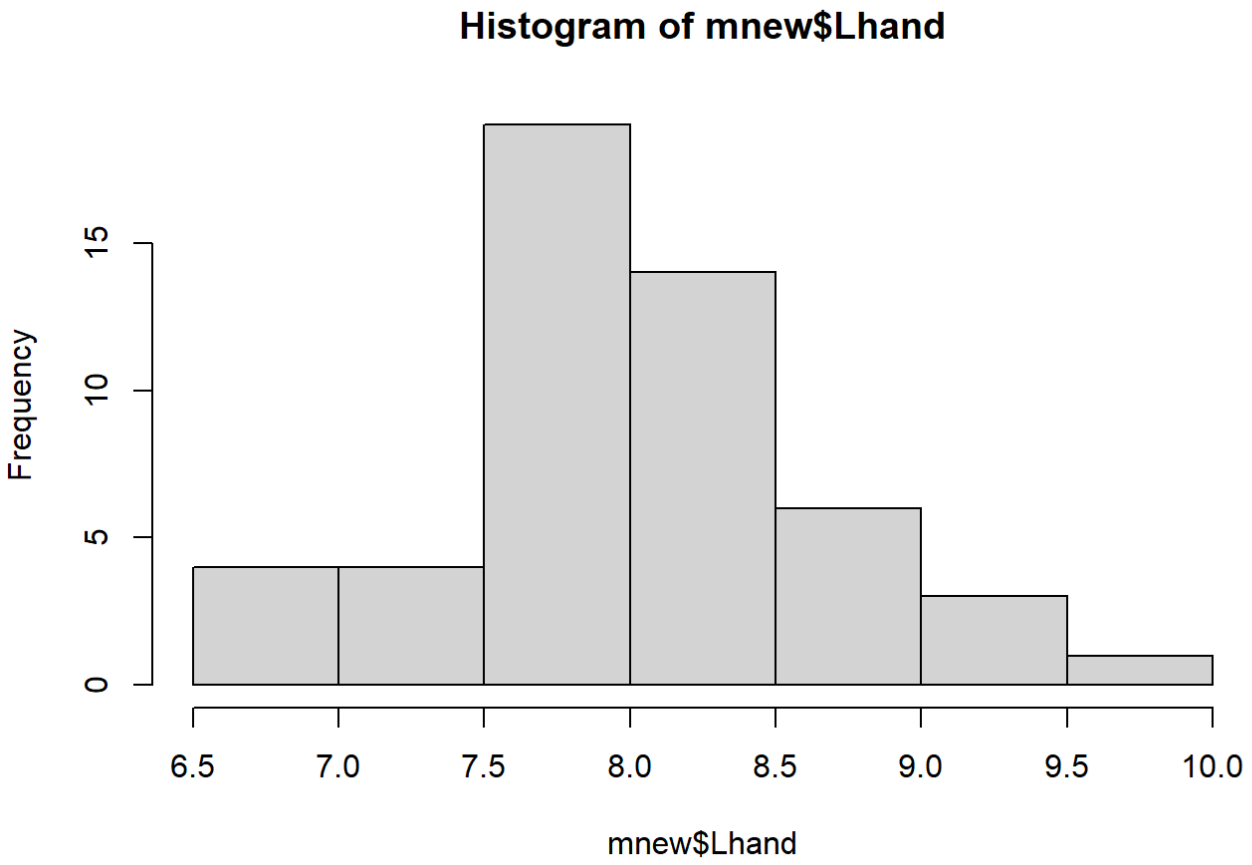

```
hist(mnew$LFoot)
```



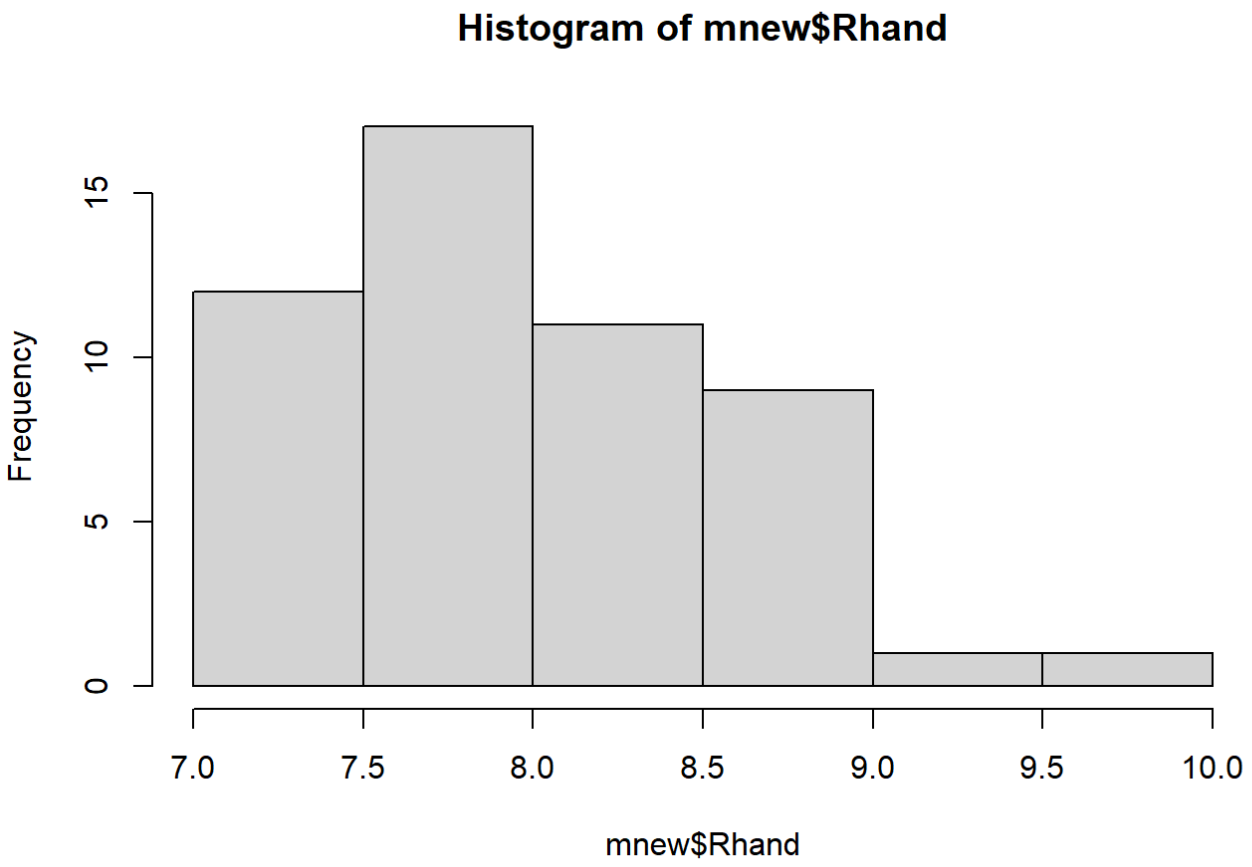

```
hist(mnew$RFoot)
```



```
hist(mnew$Lhand)
```




```
hist(mnew$Rhand)
```

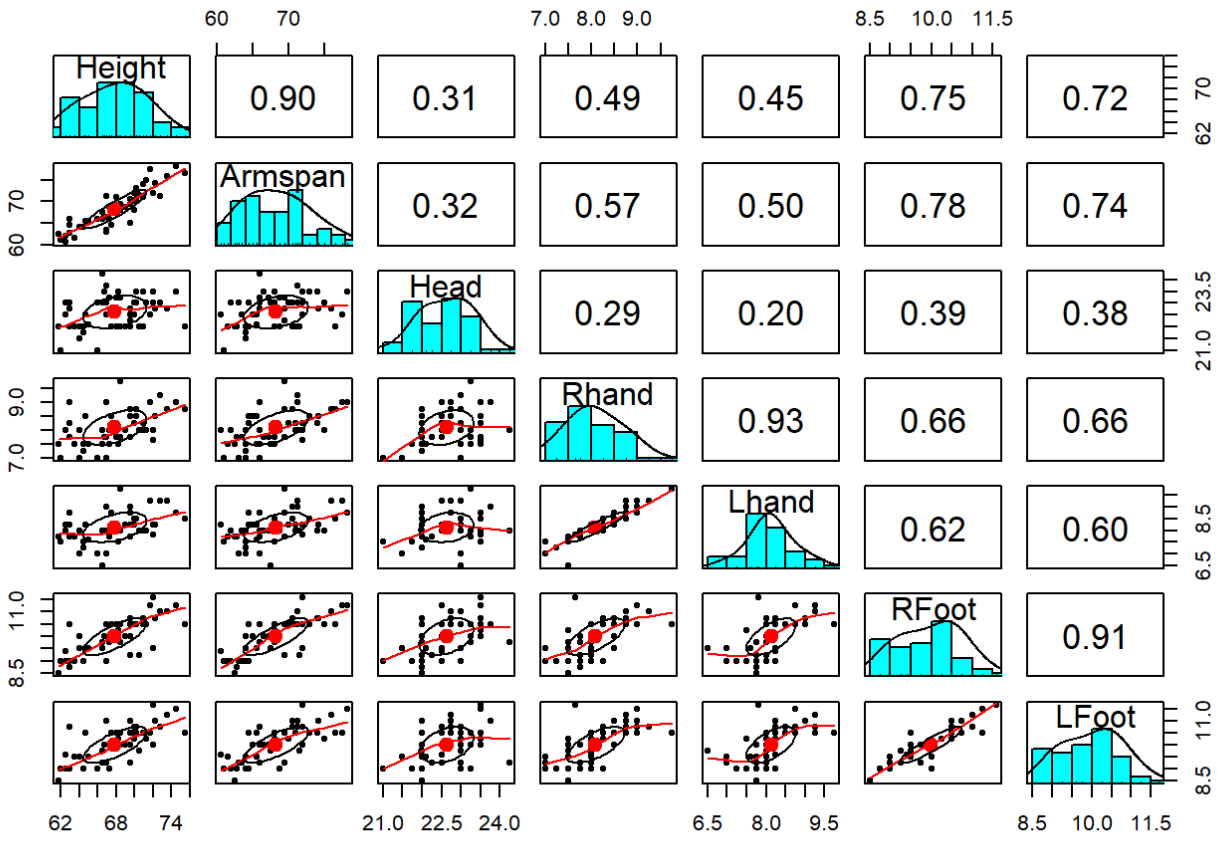



```
#  Something new.  Install the package psych, and try this:
library(psych)
```

```
##
## Attaching package: 'psych'
```

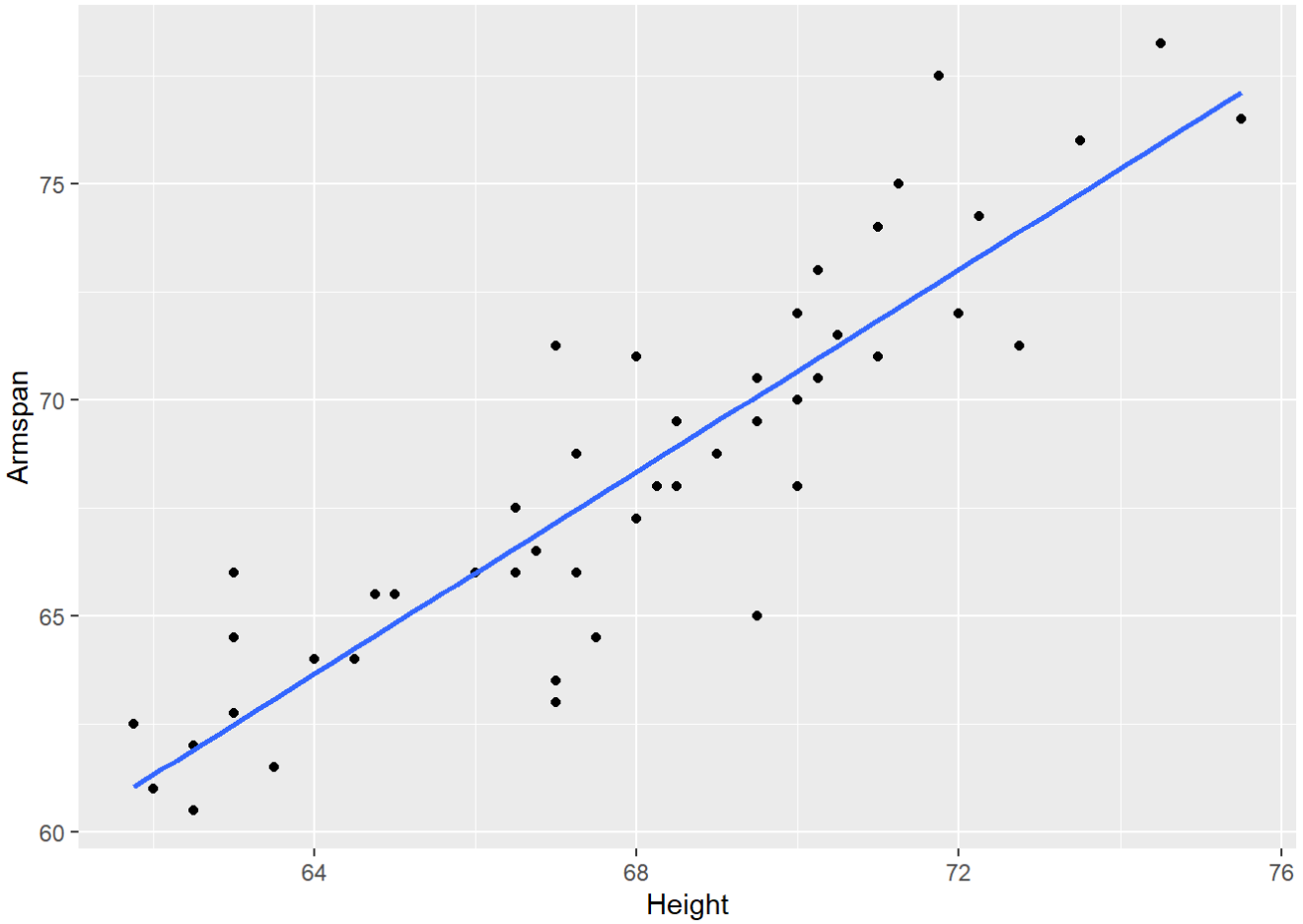
```
## The following objects are masked from 'package:ggplot2':
##
##     %+%, alpha
```

```
pairs.panels(mnew[,3:9])
```

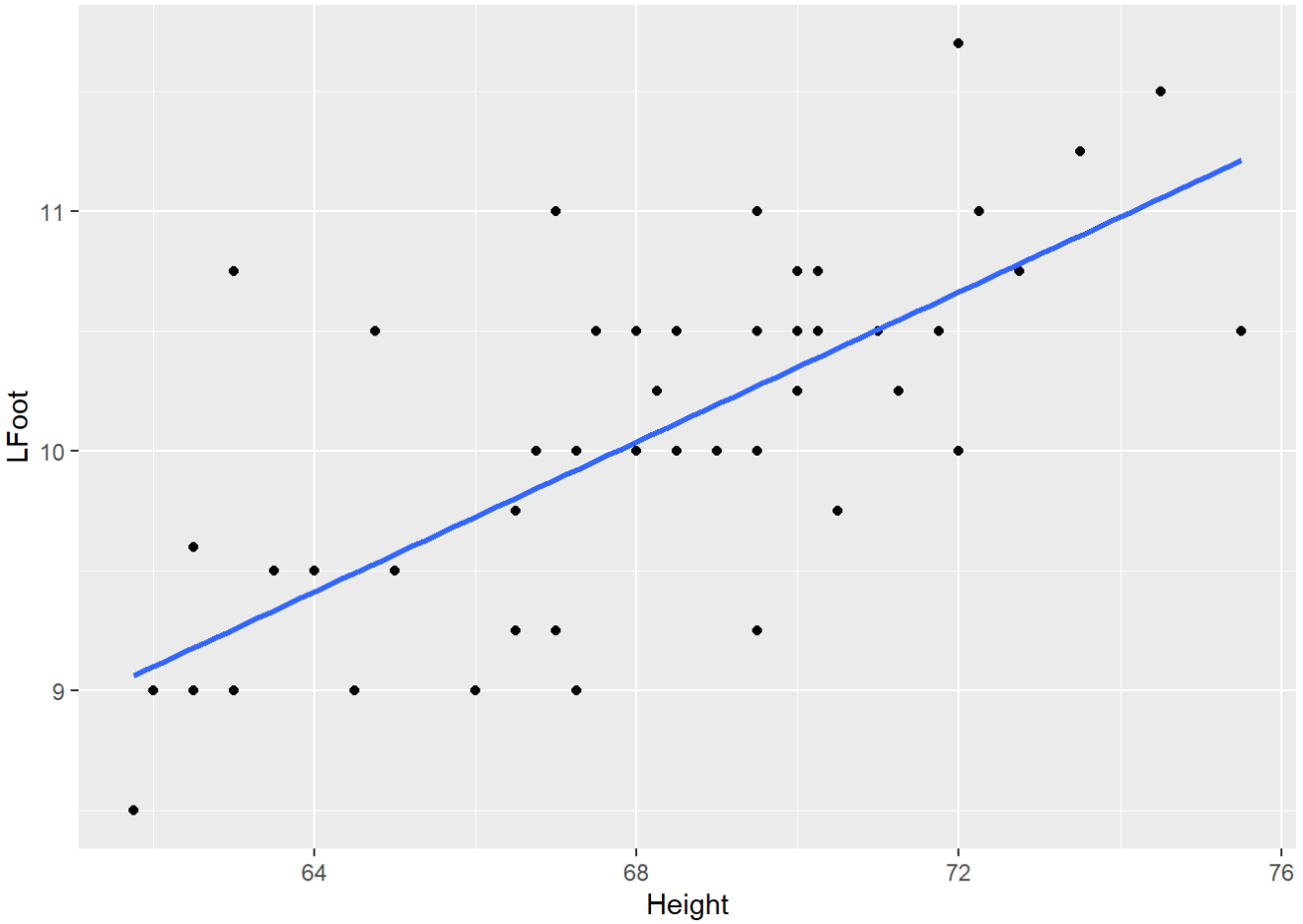


part b: Plots with regression lines: 2 pts

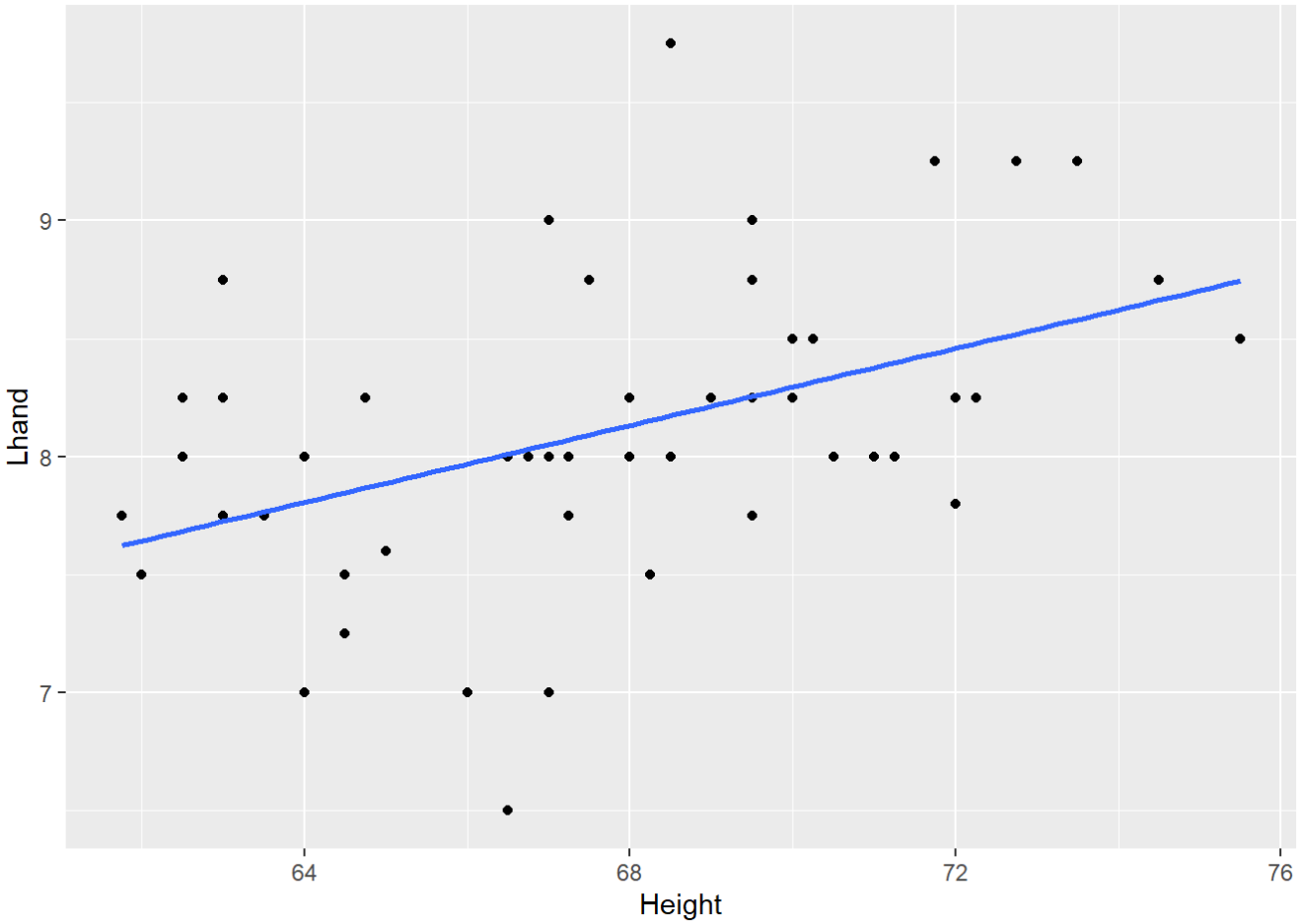
```
# Using ggplot, create a scatterplot of Armspan by Height, adding a regression line to the plot.  Do remove the confidence band on the regression line.  Repeat with LFoot by Height, then with Lhand by Height.

# Scatterplot of Armspan by Height
ggplot(data = mnew,
       mapping = aes( x = Height, y = Armspan))+
  geom_point() +
  geom_smooth(method = 'lm', se = FALSE)
```




```
# Scatterplot of LFoot by Height
ggplot(data = mnew,
       mapping = aes( x = Height, y = LFoot))+
  geom_point() +
  geom_smooth(method = 'lm', se = FALSE)
```




```
# Scatterplot of Lhand by Height
ggplot(data = mnew,
       mapping = aes( x = Height, y = Lhand))+
  geom_point() +
  geom_smooth(method = 'lm', se = FALSE)
```



part c: Linear regression equations: 2 pts

```
# Create the linear model object for Armspan by Height, as we did in class, and do
a summary of the object.   Repeat with LFoot by Height, then Lhand by Height.

# Linear model object for Armspan by Height
armspanHeight <- lm(Armspan ~ Height, data = mnew)
summary(armspanHeight)
```

```
##
## Call:
## lm(formula = Armspan ~ Height, data = mnew)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.1008 -0.8929 -0.2588  1.0874  4.7703
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -11.10286    5.48591  -2.024   0.0485 *
## Height        1.16840    0.08068  14.483   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.982 on 49 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.8106, Adjusted R-squared:  0.8068
## F-statistic: 209.7 on 1 and 49 DF,  p-value: < 2.2e-16
```

```
# Linear model object for LFoot by Height
lfootHeight <- lm(LFoot ~ Height, data = mnew)
summary(lfootHeight)
```

```
##
## Call:
## lm(formula = LFoot ~ Height, data = mnew)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.02369 -0.28542 -0.03157  0.32390  1.49208
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.58715    1.47358  -0.398    0.692
## Height       0.15627    0.02167   7.211 3.13e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5325 on 49 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.5149, Adjusted R-squared:  0.505
## F-statistic:    52 on 1 and 49 DF,  p-value: 3.125e-09
```

```
# Linear model object for Lhand by Height
lhandHeight <- lm(Lhand ~ Height, data = mnew)
summary(lhandHeight)
```

```
##
## Call:
## lm(formula = Lhand ~ Height, data = mnew)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.51043 -0.32899 -0.03081  0.25995  1.57657
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.59060    1.57559   1.644  0.10653
## Height       0.08150    0.02317   3.517  0.00095 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5694 on 49 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.2016, Adjusted R-squared:  0.1853
## F-statistic: 12.37 on 1 and 49 DF,  p-value: 0.0009504
```

part d: Interpret slopes: 2 pts

Note the slope of each regression line, and summarize each slope, as we did in class: (For each….) Write your answers in text in the white part right here:

All slopes are linear and positive.

For each additional 4 inches of height, there is a 1.16840 increase in predicted inches of Armspan.

For each additional 4 inches of height, there is a 0.15627 increase in predicted inches of left foot length.

For each additional 4 inches of height, there is a 0.08150 increase in predicted inches of left hand length.

part e: Find correlation coefficients: 1 pt

```
# Using dplyr, find the correlation coefficients for each relationship, using as mu
ch data as you can (i.e., don't remove values that are missing for variables other
than the two under consideration).

# Correlation coefficent for for Armspan by Height
ah <- mnew %>%filter(!is.na(Armspan), !is.na(Height))
cor(ah$Armspan, ah$Height)
```

```
## [1] 0.900346
```

```
# Correlation coefficent for for Armspan by Height
fh <- mnew %>%filter(!is.na(LFoot), !is.na(Height))
cor(fh$LFoot, fh$Height)
```

```
## [1] 0.7175355
```

```
# Correlation coefficent for for Armspan by Height
hh <- mnew %>%filter(!is.na(Lhand), !is.na(Height))
cor(hh$Lhand, hh$Height)
```

```
## [1] 0.4489903
```

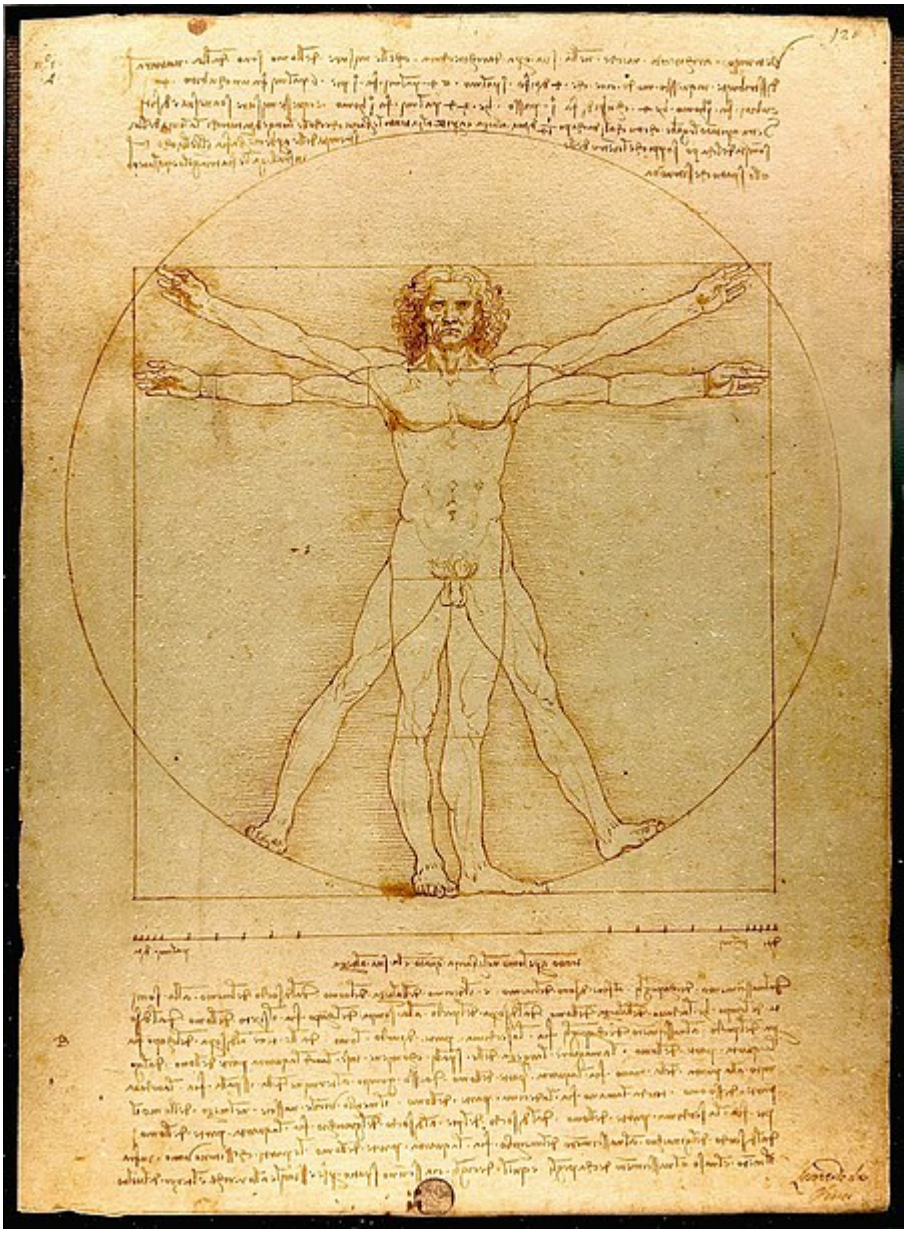
part f: Interpret 2 pts

Summarize your findings, comparing the results for the three relationships. What was the general phenomenon responsible for the relationships? Which relationship is strongest? weakest? Why would that be, in terms of what you know about human bodies?

All relationships are positive. Armspan by Height has clearly the strongest relationship (r = 0.900346), then left foot by Height(r = 0.7175355) and the least is Left Hand by Height (r = 0.4489903). I believe that the reason for these relationships is due to evolution and how our bodies were made to support us, so we are able to move bipedally.

The most obvious causation for a strong relationship between feet length and height, is that the feet are what supports our body and it makes sense that the taller you are the longer your feet should be. I think hands are not that important height wise and maybe that is why it has the weakest relationship.

Trying to think of a reason for the relationship of armspan and height, the image of the “Vitruvian Man” by Leonardo da Vinci popped up. That drawing represents Leonardo’s concept of the ideal human body proportions, and we can see how a square and a circumference surround the body. If we take a closer look, we can see that the height and the armspan perpendicular to the body, both are equal to the side length of the square. The causation of this could be so that our bodies are proportional, meaning their relative magnitudes are in balance and make sense the way they are.



Leonardo da Vinci’s Vitruvian Man