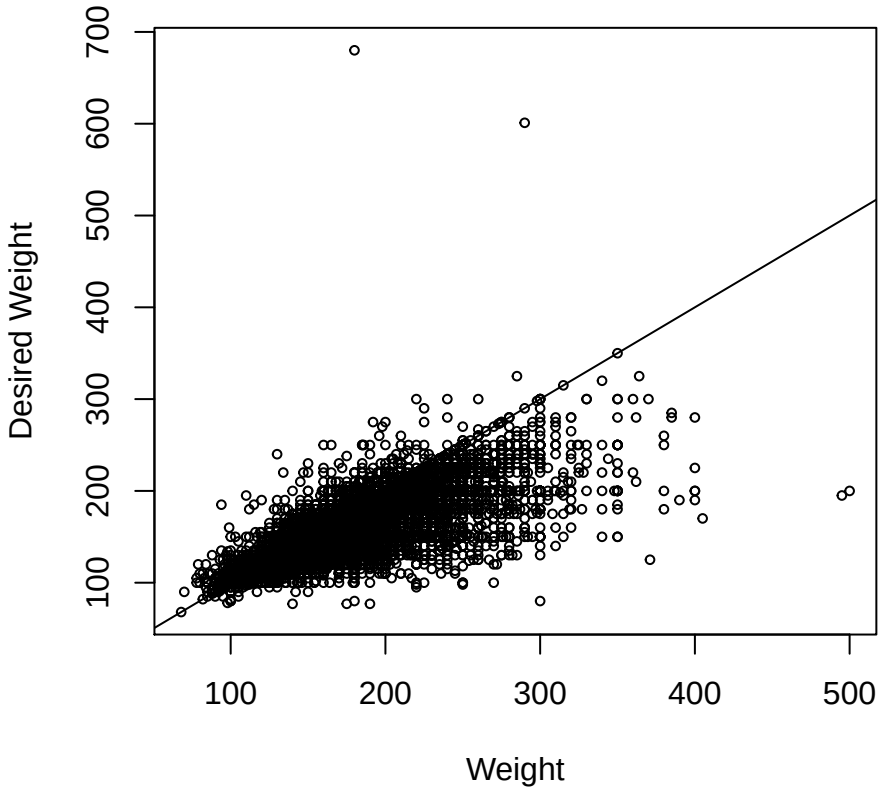
Lab 2 Solutions, STAT 630

Exercise 1

```
# first load data
data_url <- "https://github.com/ericwfox/stat630data/raw/master/cdc.csv"
cdc <- read.csv(data_url, header = TRUE)

plot(cdc$weight, cdc$wtdesire, xlab="Weight", ylab="Desired Weight", cex=0.6)
abline(0,1)
```



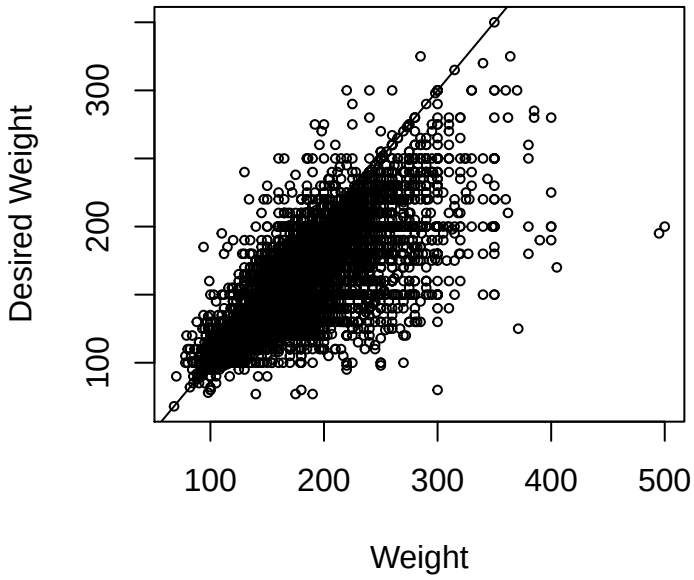
Exercise 2

The actual weights of the two individuals are 290 and 180 pounds.

```
subset(cdc, wtdesire > 500)
```

```
##         genhlth exerany hlthplan smoke100 height weight wtdesire age
## 10034 very good       1        1        1     73    290      601  56
## 16874      good       0        1        0     69    180      680  24
##       gender
## 10034      m
## 16874      m
```

```
cdc2 <- subset(cdc, wtdesire < 500)
plot(cdc2$weight, cdc2$wtdesire, xlab="Weight", ylab="Desired Weight", cex=0.6)
abline(0,1)
```



Exercise 3

```
cdc_m_ex <- subset(cdc, exerany==1 & gender == "m")
summary(cdc_m_ex$weight)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    94.0   165.0   185.0   188.9   206.0   500.0
```

```
summary(cdc_m_ex$wtdesire)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    77.0   162.0   175.0   179.2   190.0   601.0
```

Exercise 4

```
addmargins(table(cdc$genhlth, cdc$exerany))
```

```
##
##                0     1   Sum
##   excellent   762  3895  4657
##   fair        857  1162  2019
##   good       1731  3944  5675
##   poor        384   293   677
##   very good  1352  5620  6972
##   Sum        5086 14914 20000
```

What proportion of respondents that reported to be in excellent health exercised in the past month?

```
3895/4657
```

```
## [1] 0.8363753
```

Note: this can be written as the conditional probability $P(\text{exercise}|\text{excellent health})$.

What proportion of respondents that reported to be in poor health exercised in the past month?

```
293/677
```

```
## [1] 0.4327917
```

Note: this can be written as the conditional probability $P(\text{exercise}|\text{poor health})$.