FinalProject

```
MovieData = read_csv("MovieData - MovieData (3).csv")
MovieData
```

```
## # A tibble: 1,936 x 14
##    movie_name production_year movie_odid production_budg~ domestic_box_of~
##    <chr>                <dbl>      <dbl>            <dbl>            <dbl>
##  1 Madea's F~            2006    8220100         10000000         63257940
##  2 Krrish               2006   58540100         10000000          1430721
##  3 End of th~           2006   34620100         10000000         11748661
##  4 A Prairie~           2006   24910100         10000000         20342852
##  5 Saw III              2006    5840100         10000000         80238724
##  6 Employee ~           2006   19540100         10000000         28444855
##  7 Edmond               2006   78520100         10000000           131719
##  8 Peaceful ~           2006   52640100         10000000          2893666
##  9 Turistas             2006   42200100         10000000          7027762
## 10 Kabhi Alv~           2006   51280100         10750000          3275443
## # ... with 1,926 more rows, and 9 more variables:
## #   international_box_office <dbl>, rating <chr>, creative_type <chr>,
## #   source <chr>, production_method <chr>, genre <chr>, sequel <dbl>,
## #   running_time <dbl>, marvel <chr>
```

```
SuperHero_MovieData <- filter(MovieData, creative_type == "Super Hero")
SuperHero_MovieData
```

```
## # A tibble: 60 x 14
##    movie_name production_year movie_odid production_budg~ domestic_box_of~
##    <chr>                <dbl>      <dbl>            <dbl>            <dbl>
##  1 My Super ~           2006   23230100         30000000         22530295
##  2 Zoom                 2006   34430100         35000000         11989328
##  3 X-Men: Th~           2006     620100        210000000        234362462
##  4 Superman ~           2006     940100        232000000        200120000
##  5 Ghost Rid~           2007    3310100        120000000        115802596
##  6 Fantastic~           2007    2500100        120000000        131921738
##  7 Spider-Ma~           2007     170100        258000000        336530303
##  8 Superhero~           2008   20910100         35000000         26638520
##  9 Punisher:~           2008   40290100         35000000          8050977
## 10 The Spirit           2008   25460100         60000000         19806188
## # ... with 50 more rows, and 9 more variables:
## #   international_box_office <dbl>, rating <chr>, creative_type <chr>,
## #   source <chr>, production_method <chr>, genre <chr>, sequel <dbl>,
## #   running_time <dbl>, marvel <chr>
```

```
Action_MovieData <- filter(MovieData, genre == "Action")
Action_MovieData
```

```
## # A tibble: 311 x 14
##    movie_name production_year movie_odid production_budg~ domestic_box_of~
##    <chr>                <dbl>      <dbl>            <dbl>            <dbl>
##  1 Krrish               2006   58540100         10000000          1430721
##  2 Crank                2006   19850100         12000000         27838408
##  3 The Marine           2006   26360100         15000000         18844784
##  4 Running S~           2006   42740100         17000000          6855137
```

```
##  5 Renaissan~             2006    83460100           18000000             70644
##  6 BloodRayne             2006    54100100           25000000           2405420
##  7 Ultraviol~             2006    26670100           30000000          18522064
##  8 Wu ji                  2006    64990100           30000000            669625
##  9 Alex Ride~             2006    65100100           40000000            659210
## 10 Apocalypto             2006   130090100           40000000          50866635
## # ... with 301 more rows, and 9 more variables:
## #   international_box_office <dbl>, rating <chr>, creative_type <chr>,
## #   source <chr>, production_method <chr>, genre <chr>, sequel <dbl>,
## #   running_time <dbl>, marvel <chr>
```

```
Marvel_SuperHero_MovieData <- filter(MovieData, marvel == "Yes")
Marvel_SuperHero_MovieData
```

```
## # A tibble: 20 x 14
##    movie_name production_year movie_odid production_budg~ domestic_box_of~
##    <chr>                <dbl>      <dbl>            <dbl>            <dbl>
##  1 The Incre~            2008    2380100        137500000        134806913
##  2 Iron Man              2008     220100        186000000        318604126
##  3 Iron Man 2            2010  117940100        170000000        312433331
##  4 Captain A~            2011  124300100        140000000        176654505
##  5 Thor                 2011  109130100        150000000        181030624
##  6 The Aveng~            2011  117960100        225000000        623279547
##  7 Iron Man 3            2012  146800100        200000000        408992272
##  8 Thor: The~            2013  163710100        150000000        206362140
##  9 Captain A~            2014  174890100        170000000        259746958
## 10 Guardians~            2014  165370100        170000000        333172112
## 11 Avengers:~            2014  179490100        330600000        459005868
## 12 Ant-Man              2015  133110100        130000000        180202163
## 13 Captain A~            2015  202820100        250000000        408084349
## 14 Doctor St~            2016  218470100        165000000        232641920
## 15 Spider-Ma~            2016  224140100        175000000        334201140
## 16 Guardians~            2016  211440100        200000000        389813101
## 17 Ant-Man a~            2017  243840100        130000000        216648740
## 18 Thor: Rag~            2017  218480100        180000000        315058289
## 19 Black Pan~            2017  140380100        200000000        700059566
## 20 Avengers:~            2017  202830100        300000000        678815482
## # ... with 9 more variables: international_box_office <dbl>, rating <chr>,
## #   creative_type <chr>, source <chr>, production_method <chr>,
## #   genre <chr>, sequel <dbl>, running_time <dbl>, marvel <chr>
```
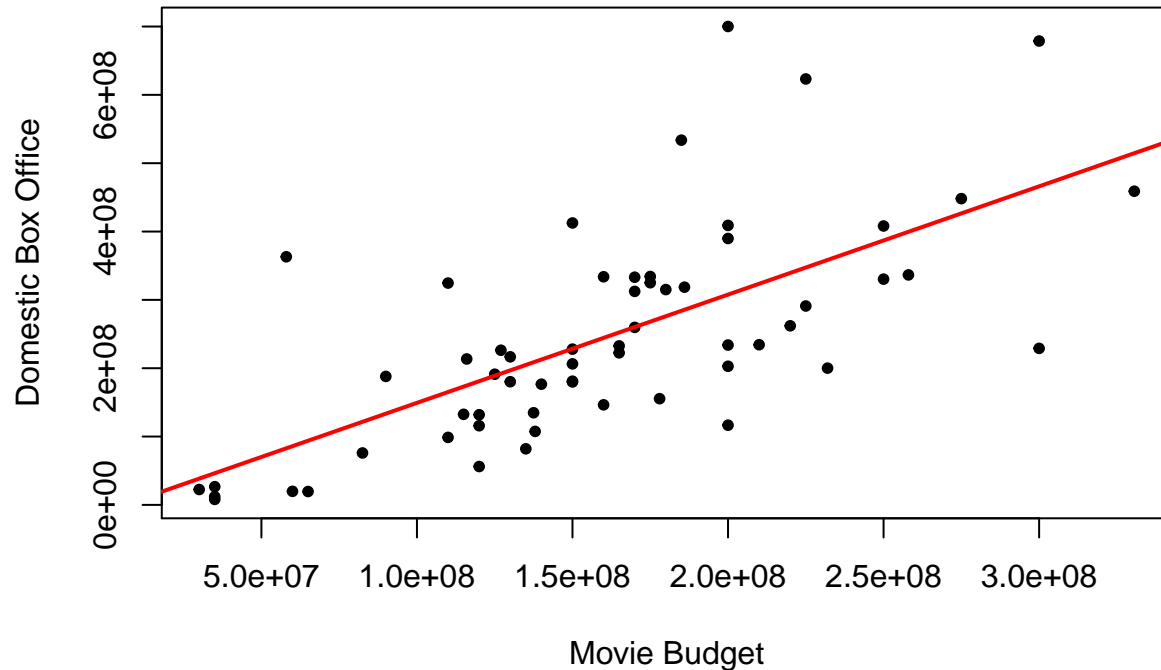
Super Hero Movies

```
budget <- SuperHero_MovieData$production_budget
domestic <- SuperHero_MovieData$domestic_box_office
international <- SuperHero_MovieData$international_box_office
plot(budget,domestic,
     main = "Domestic Budget vs Box office for super hero movies",
     xlab = "Movie Budget",
     ylab = "Domestic Box Office",
     pch = 20)
abline(lm(domestic~budget),
```

```
        col="red",
        lwd = 2)
```

Domestic Budget vs Box office for super hero movies


```r
cor(budget,domestic)
```

```
## [1] 0.6746
```

```r
cor(budget,international)
```

```
## [1] 0.7007
```

```r
lm(domestic~budget)
```

```
##
## Call:
## lm(formula = domestic ~ budget)
##
## Coefficients:
## (Intercept)        budget
##    -9.09e+06      1.58e+00
```

```r
SuperHero_lm <- lm(international~budget)
summary(SuperHero_lm)
```

```
##
## Call:
## lm(formula = international ~ budget)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -4.00e+08  -1.06e+08  -2.07e+05   6.90e+07   6.08e+08
##
```

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -6.75e+07   6.42e+07   -1.05      0.3
## budget       2.76e+00   3.69e-01    7.48  4.6e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.9e+08 on 58 degrees of freedom
## Multiple R-squared:  0.491,  Adjusted R-squared:  0.482
## F-statistic:   56 on 1 and 58 DF,  p-value: 4.57e-10
```
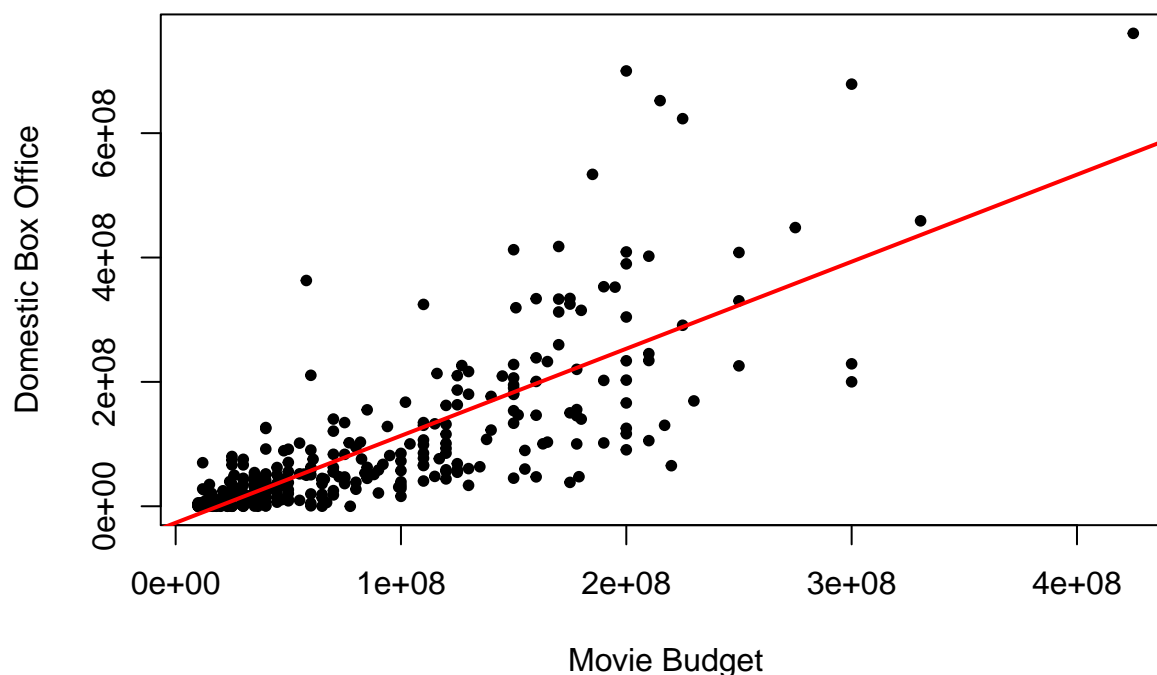
```r
qt(.025, df = 58)
```

```
## [1] -2.002
```

```r
2.76e+00 + c(-2,2)*3.69e-01
```

```
## [1] 2.022 3.498
```

Action Movie Data

```r
budget <- Action_MovieData$production_budget
domestic <- Action_MovieData$domestic_box_office
international <- Action_MovieData$international_box_office
plot(budget,domestic,
     main = "Domestic Budget vs Box office for action movies",
     xlab = "Movie Budget",
     ylab = "Domestic Box Office",
     pch = 20)
abline(lm(domestic~budget),
       col="red",
       lwd = 2)
```

Domestic Budget vs Box office for action movies


```r
cor(budget,domestic)
```

```
## [1] 0.7788
```

```r
Action_lm <- lm(domestic~budget)
lm(international~budget)
```

```
##
## Call:
## lm(formula = international ~ budget)
##
## Coefficients:
## (Intercept)        budget
##   -6.71e+07      2.84e+00
```

```r
summary(Action_lm)
```

```
##
## Call:
## lm(formula = domestic ~ budget)
##
## Residuals:
##       Min       1Q    Median       3Q       Max
## -2.16e+08 -3.48e+07 -5.25e+05  2.08e+07  4.47e+08
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.66e+07   7.24e+06   -3.67  0.00029 ***
## budget       1.40e+00   6.41e-02   21.82  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 79700000 on 309 degrees of freedom
## Multiple R-squared:  0.607,   Adjusted R-squared:  0.605
## F-statistic:  476 on 1 and 309 DF,  p-value: <2e-16
```

```r
qt(.05, df = 309)
```
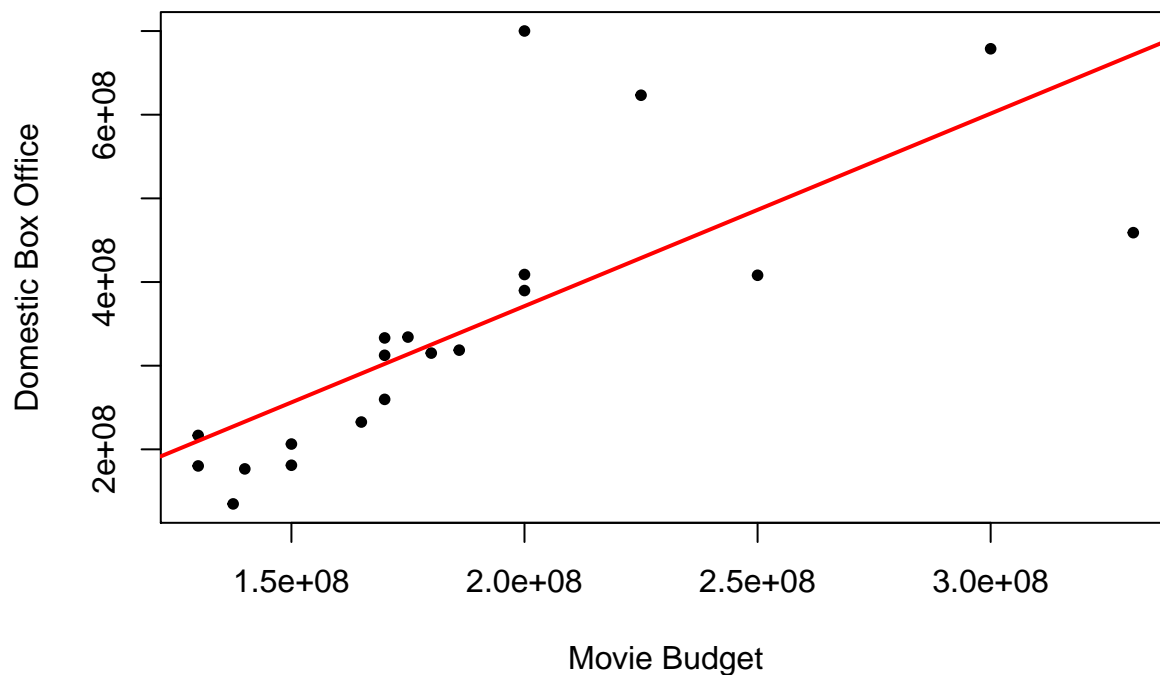
```
## [1] -1.65
```

```r
1.40e+00 + c(-2,2)*6.41e-02
```

```
## [1] 1.272 1.528
```

Marvel Movie Data

```r
budget <- Marvel_SuperHero_MovieData$production_budget
domestic <- Marvel_SuperHero_MovieData$domestic_box_office
international <- Marvel_SuperHero_MovieData$international_box_office
plot(budget,domestic,
     main = "Domestic Budget vs Box office for marvel movies",
     xlab = "Movie Budget",
     ylab = "Domestic Box Office",
     pch = 20)
abline(lm(domestic~budget),
       col="red",
       lwd = 2)
```

Domestic Budget vs Box office for marvel movies


```r
cor(budget,domestic)
```

```
## [1] 0.7445
```

```
Marvel_lm <-lm(domestic~budget)
summary(Marvel_lm)
```

```
##
## Call:
## lm(formula = domestic ~ budget)
##
## Residuals:
##       Min       1Q    Median        3Q       Max
## -2.13e+08 -5.68e+07 -1.52e+07  2.32e+07  3.29e+08
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -8.91e+07   9.49e+07   -0.94  0.36028
## budget       2.30e+00   4.86e-01    4.73  0.00017 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.14e+08 on 18 degrees of freedom
## Multiple R-squared:  0.554,  Adjusted R-squared:  0.53
## F-statistic: 22.4 on 1 and 18 DF,  p-value: 0.000167
```

```
qt(.025, df = 18)
```

```
## [1] -2.101
```

```
2.30e+00 + c(-2.1,2.1)*6.41e-02
```

```
## [1] 2.165 2.435
```