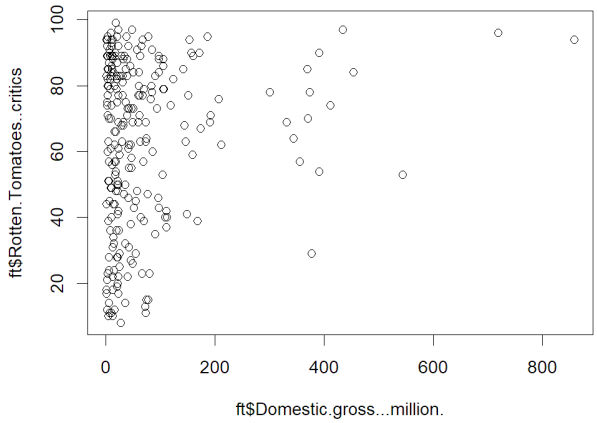
```
ft <- read.csv("pathhere")

# Part 1 Naïve Inplementation
modelDom <-
lm(Domestic.gross...million.~Rotten.Tomatoes..critics+Metacritic..critics+Rotten.Tomatoe
s.Audience+Metacritic.Audience+Opening.weekend...million.+Foreign.Gross...million.+W
orldwide.Gross...million.+Budget...million.+IMDb.Rating,data = ft)
summary(modelDom)
```

```
Coefficients:
                            Estimate Std. Error  t value Pr(>|t|)
(Intercept)                -0.357527   1.345345   -0.266    0.791
Rotten.Tomatoes..critics   -0.002774   0.016454   -0.169    0.866
Metacritic..critics        -0.010310   0.026104   -0.395    0.693
Rotten.Tomatoes.Audience   -0.008239   0.012706   -0.648    0.517
Metacritic.Audience         0.004147   0.014339    0.289    0.773
Opening.weekend...million.  0.008388   0.011903    0.705    0.482
Foreign.Gross...million.   -0.996634   0.005777 -172.522   <2e-16
Worldwide.Gross...million.  0.996641   0.004621  215.658   <2e-16
Budget...million.           0.002325   0.003311    0.702    0.483
IMDb.Rating                 0.196161   0.339069    0.579    0.563
```

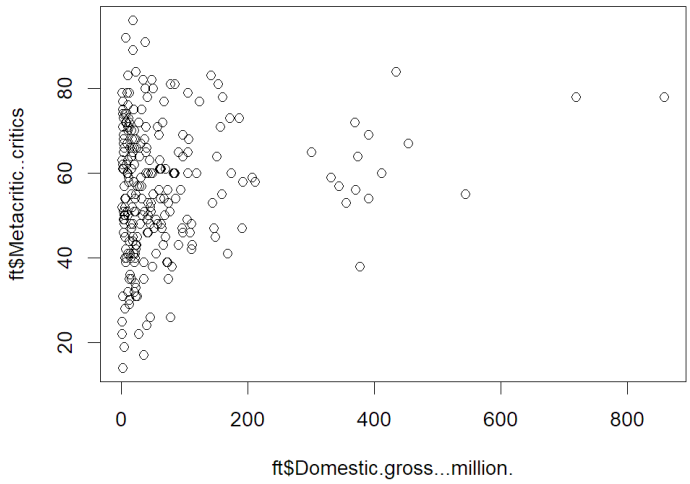
```
# Comment:
# Only Foreign.Gross...million. & Worldwide.Gross...million. correlate with
Domestic.gross...million.

# Part 2 Individual Correlation
# each variable test correlation with domestic
# (*) refers to the signif. code
# Rotten Tomato Critics *
modelRotTomCrit <- lm(Domestic.gross...million.~Rotten.Tomatoes..critics,data = ft)
summary(modelRotTomCrit)
plot(ft$Domestic.gross...million.,ft$Rotten.Tomatoes..critics)
```

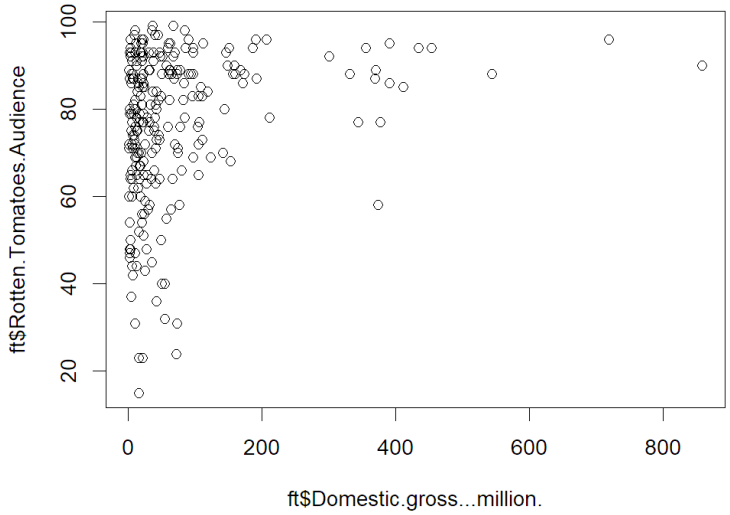



```
# Mate Critics *
modelMetaCrit <- lm(Domestic.gross...million.~Metacritic..critics,data = ft)
```

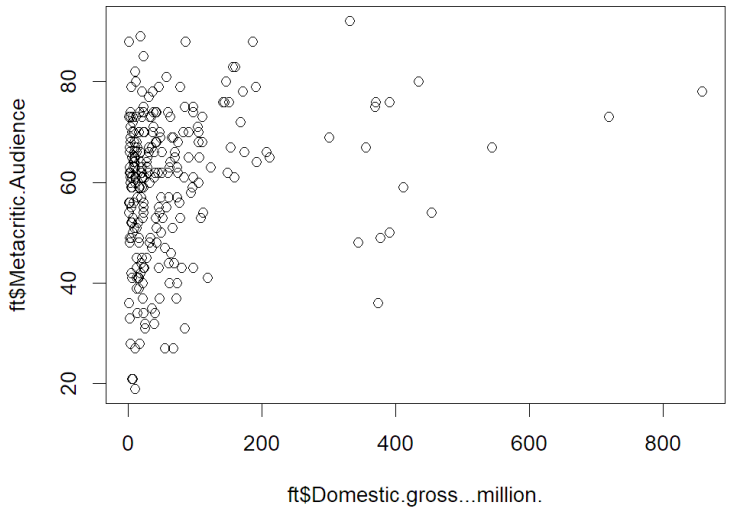
```
summary(modelMetaCrit)
plot(ft$Domestic.gross...million.,ft$Metacritic..critics)
```




```
# Rotten Tomato Audience ***
modelRotTomAudience <- lm(Domestic.gross...million.~Rotten.Tomatoes.Audience,data
= ft)
summary(modelRotTomAudience)
plot(ft$Domestic.gross...million.,ft$Rotten.Tomatoes.Audience)
```

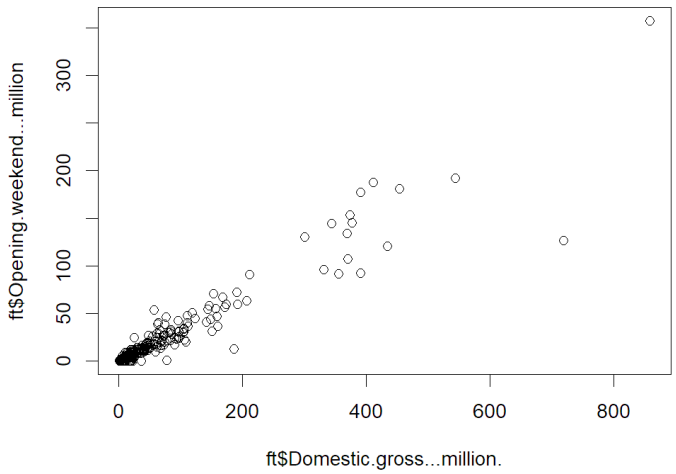



```
# Meta Audience **
modelMetaAudience <- lm(Domestic.gross...million.~Metacritic.Audience,data = ft)
summary(modelMetaAudience)
plot(ft$Domestic.gross...million.,ft$Metacritic.Audience)
```



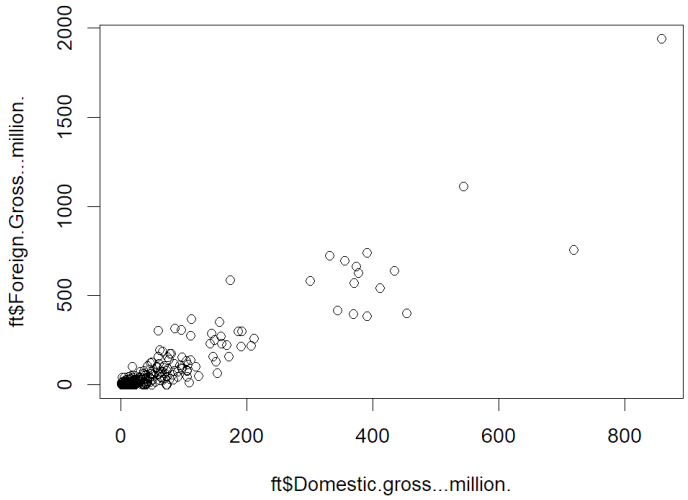
Note here: for the first 4 varibale, seems NO correlation to domestic

```
# Opening Weekend ***
modelOpen <- lm(Domestic.gross...million.~Opening.weekend...million.,data = ft)
summary(modelOpen)
plot(ft$Domestic.gross...million.,ft$Opening.weekend...million)
# Positive linear relation
```




```
#Foreigh Gross ***
modelForeigh <- lm(Domestic.gross...million.~Foreign.Gross...million.,data = ft)
summary(modelForeigh)
plot(ft$Domestic.gross...million.,ft$Foreign.Gross...million.)
```

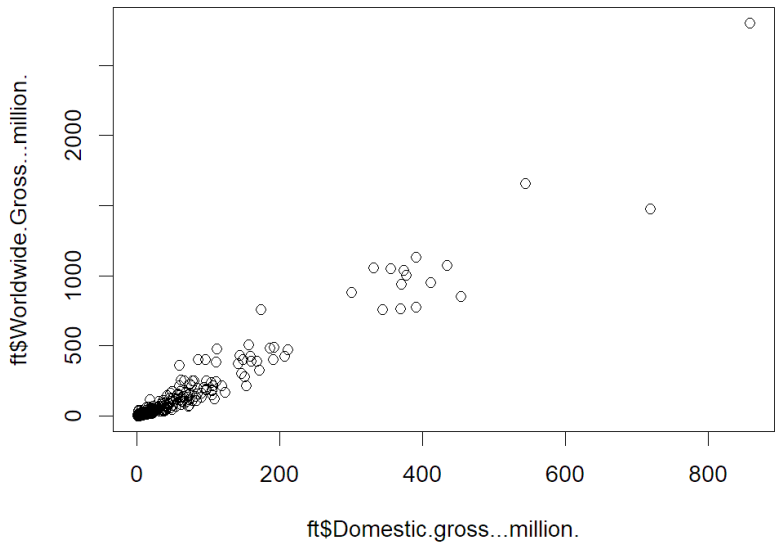
Positive linear relation, spread out as domestic increase, but no transformation could apply



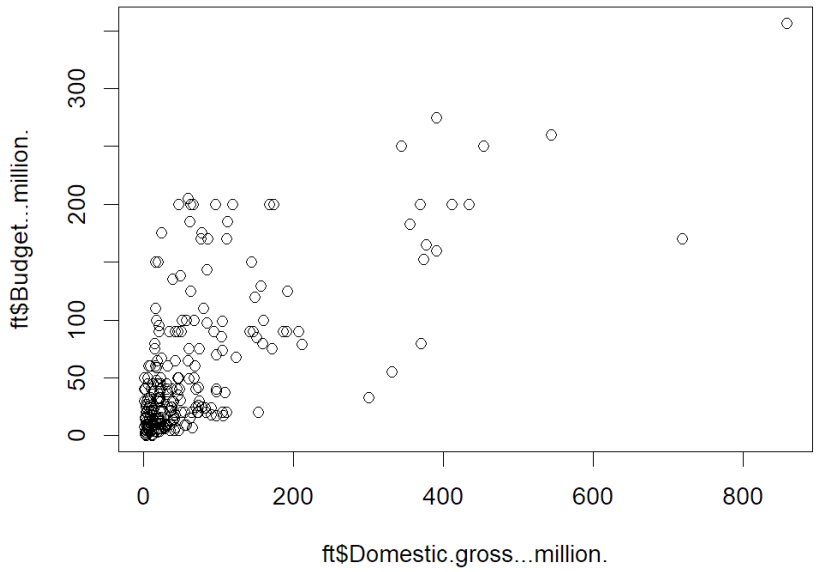
WorldWide Gross ***

```
modelWorldWide <- lm(Domestic.gross...million.~Worldwide.Gross...million.,data = ft)
summary(modelWorldWide)
plot(ft$Domestic.gross...million.,ft$Worldwide.Gross...million.)
```

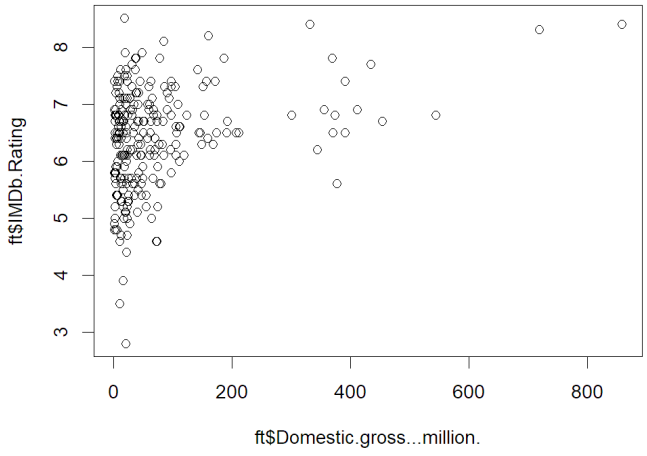
Positive linear relation, spread out as domestic increase, but no transformation could apply.



```
# Budget ***
modelBudget <- lm(Domestic.gross...million.~Budget...million.,data = ft)
summary(modelBudget)
plot(ft$Domestic.gross...million.,ft$Budget...million.)
# Normally Positive linear relation, spread out as domestic increase
# if apply log(y) transformation, may not be better
```




```
# IMDB ***
modelIMDB <- lm(Domestic.gross...million.~IMDb.Rating,data = ft)
summary(modelIMDB)
plot(ft$Domestic.gross...million.,ft$IMDb.Rating)
# Positive linear relation, spread out as domestic increase, but no transformation could
apply
```



Best Model choice

If we combine all any other variable other than just Foreign and World Wide, other variable will not be significant

```
modelOnlySign <- lm(Domestic.gross...million.~
                      +Foreign.Gross...million.+Worldwide.Gross...million.,data = ft)
summary(modelOnlySign)
```

```
Coefficients:
                           Estimate Std. Error  t value Pr(>|t|)
(Intercept)               -0.151978   0.166837   -0.911    0.363
Foreign.Gross...million.  -0.997673   0.004917 -202.894   <2e-16 ***
Worldwide.Gross...million.  0.998687   0.003207  311.416   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.196 on 247 degrees of freedom
Multiple R-squared:  0.9996,    Adjusted R-squared:  0.9996
F-statistic: 3.276e+05 on 2 and 247 DF,  p-value: < 2.2e-16
```

Note here, If we research on Worldwide Gross, it’s similar to domestic gross