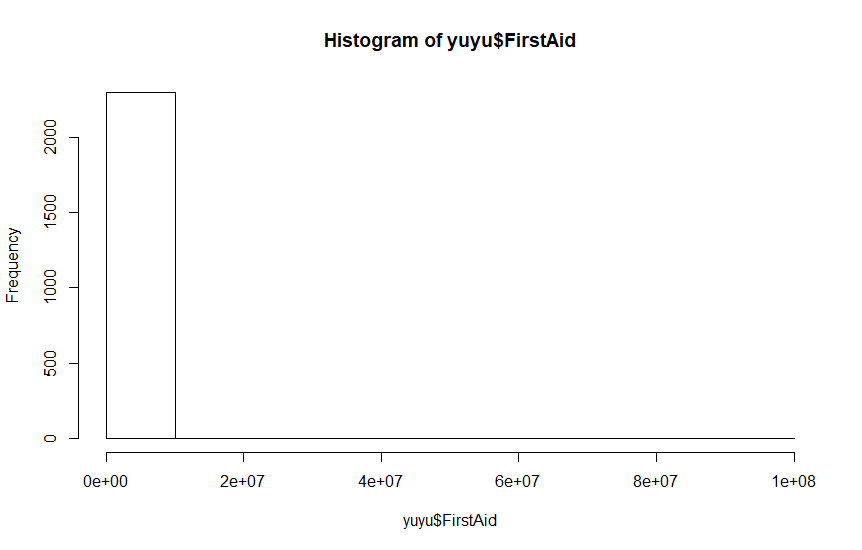
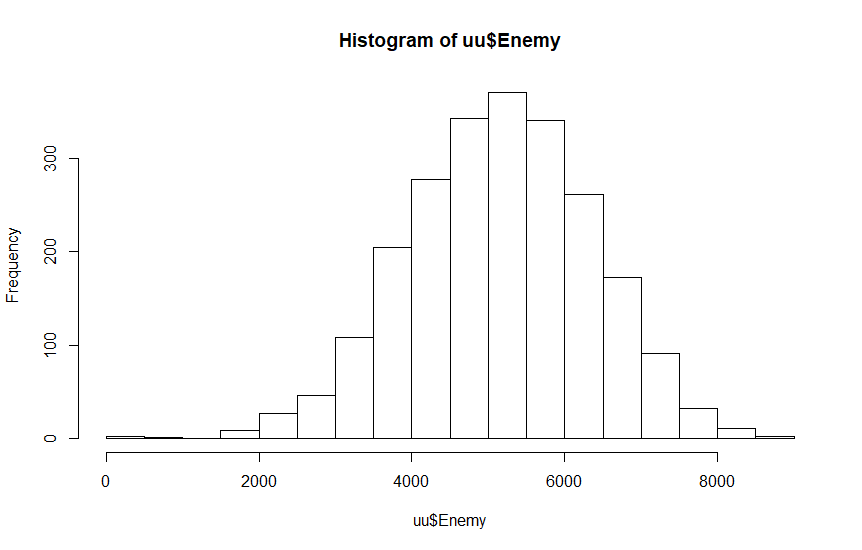
How to win our war

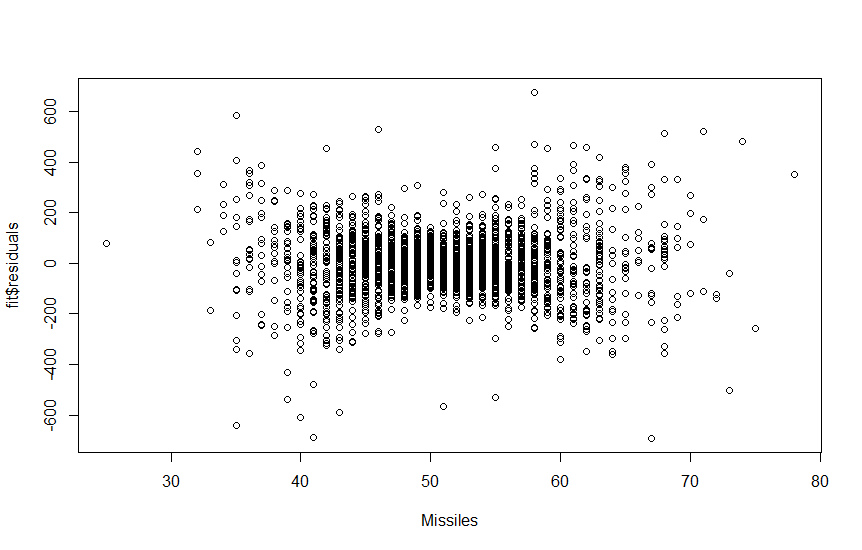
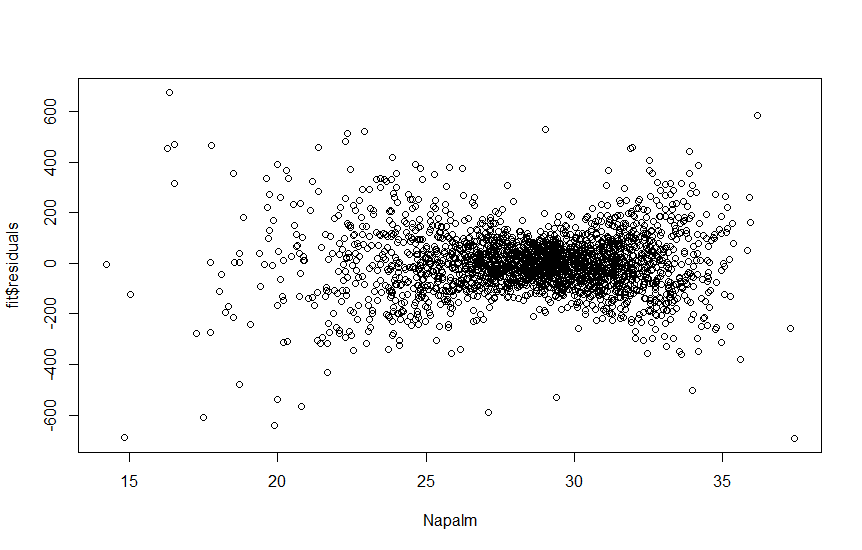
With an analysis of the existing data, we would be able to know how other factors including stock, firepower, first aid, and media and so on will influence the number of enemy casualties and civilian casualties. Based on this relationship, we may be able to predict the number of enemy casualties and civilian casualties according to the battlefield situation, and thus adjust our strategies. What’s more, if we know how the bombs, fire powers, first aid will influence the number of enemy casualties and civilian casualties, we can take measures to maximize the number of enemy casualties, and minimize the numbers of civilian casualties. Through this way, we can improve our battle plans and increase the possibility of winning the war at the least cost.

## How to maximize the number of enemy casualties?

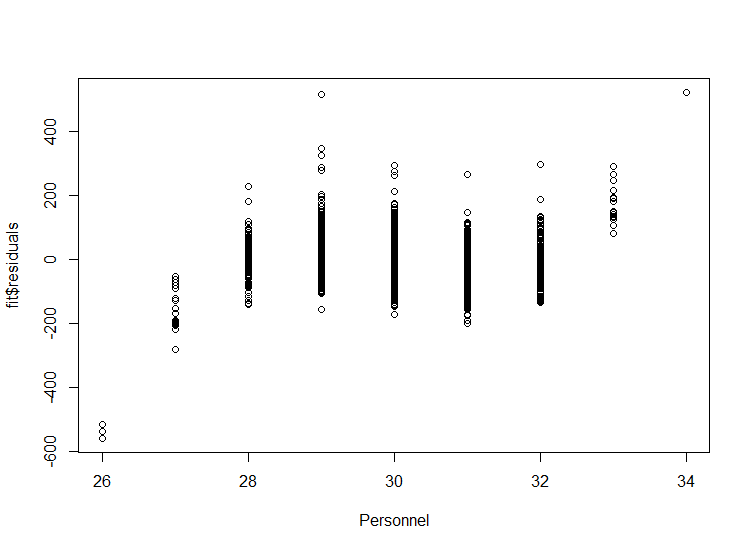
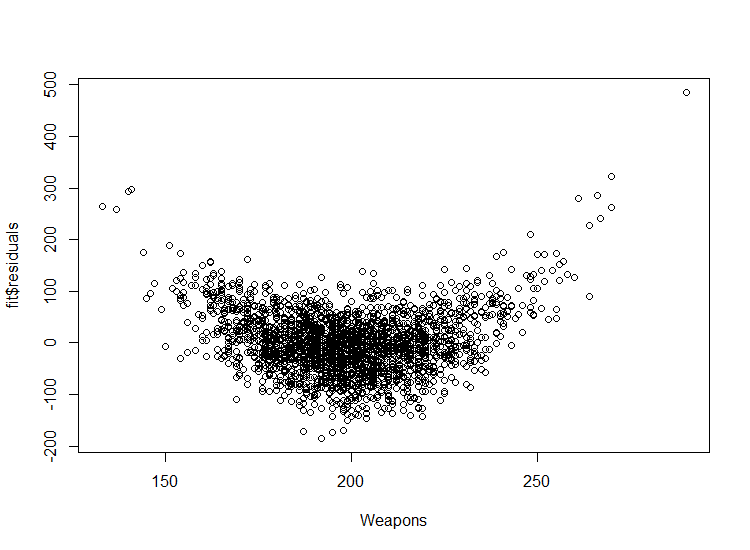


In the first part, I want to develop a model to predict the number of enemy casualties based on other factors. I used R-studio to analyze the data at an alpha of 0.05. When I look at the data, I find most time there are 4000-6000 enemy casualties. There is a plot that there is no enemy casualties, and I take it as an error and delete it. I examined all the other variables, and I find similarly there is an especially big value of FirstAid so I also deleted that data. I made a linear model with all the variables to predict the enemy causalities, but the residual plots all looks bad so I know I made a wrong model.

To fix the model, I looked at the residual plots against each variable. I find the residual plots against both Missiles and Napalm look like a bowtie shape, so I realized that they are interacting with each other. I added the numerical interaction between them and looked at the residual plots again.

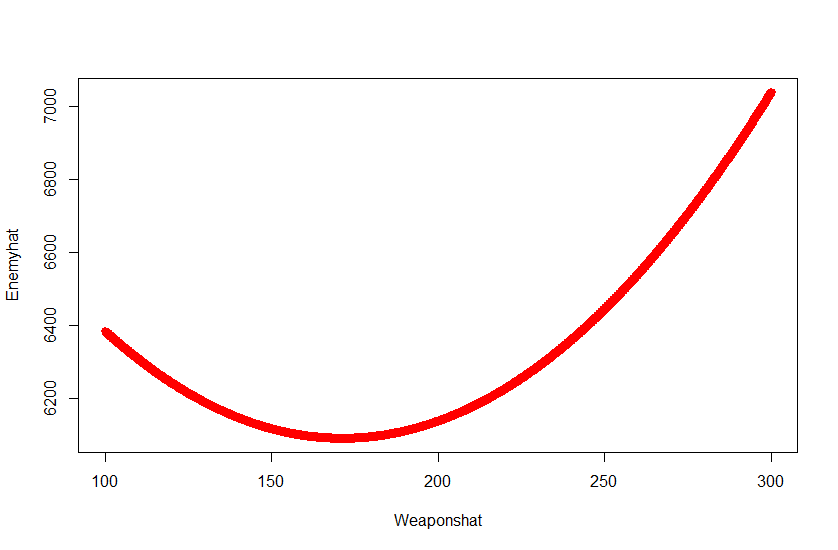
 

Then I find that the residual plot against Personnel looks like polynomial. Then I add a cubic effect to it. I also add a quadratic term to Weapons, because the residual plot looks told me to do so. Now the residual plots looks much nicer. I don’t have any logged effect in my model, because I can’t see any sign of it from the residuals plots against each variables.

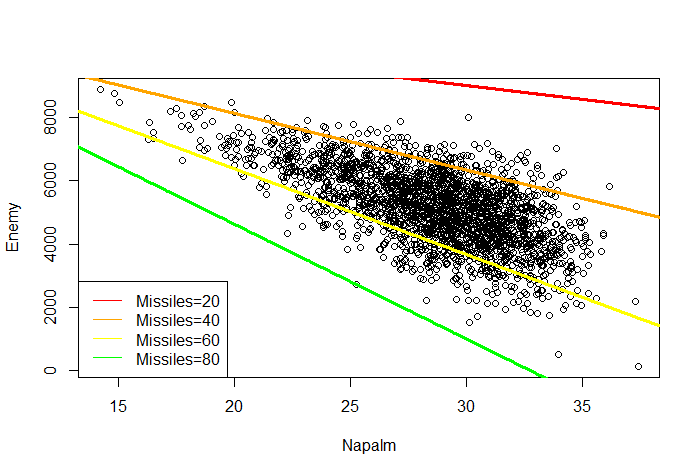
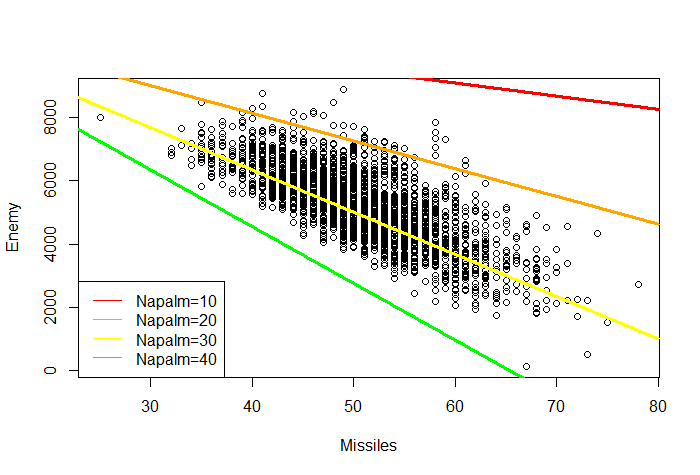
My final equation is as following:

Enemy<- -2.650522e+05-1.966965e+01\*Weapons+ 5.737243e-02 \*Weapons2+ 4.258913e+00 \*Missiles+ 4.495025e+00 \*Napalm-2.107412e+02 \*Media+ 2.804106e+04 \*Personnel-9.381782e+02 \*Personnet2+ 1.042475e+01 \*Personnel3+ 1.457744e+00\*Temperature -2.484868e+02 \*IG88Installed -4.594089e+00 \*Missiles\*Napalm

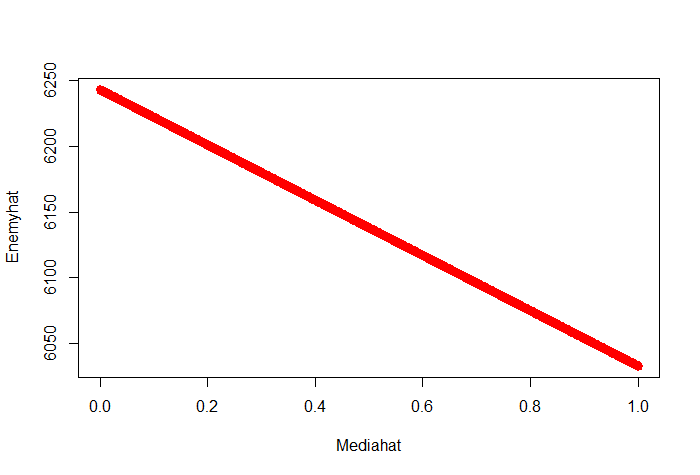


As we can see from the picture, the effect of weapons is quadratic. As the budget for Weapons increase, the number of enemy casualties decrease at first, and increase then. That’s probably because as the budget increase at first, warriors are not familiar with the weapons that are newly purchased. As budget keep increasing, warriors have more weapons to fight, the number of enemy causalities increase quickly.

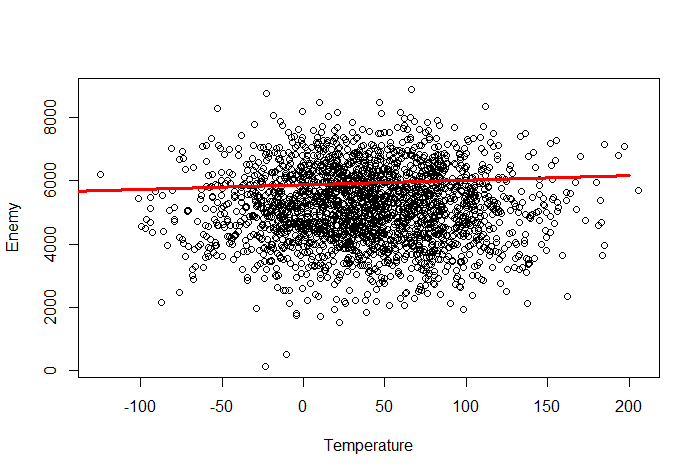
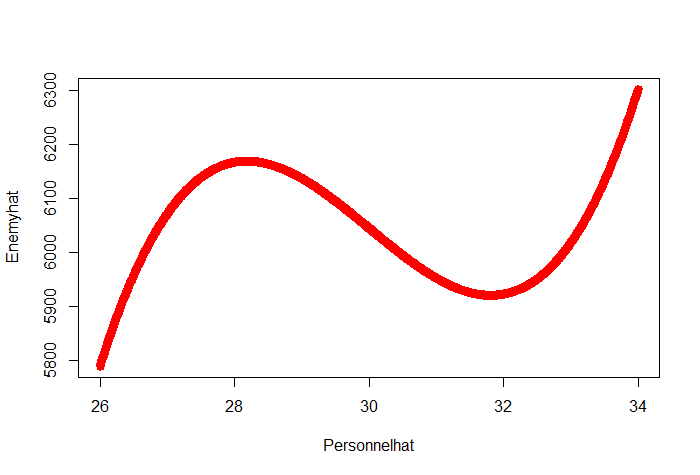
As we can see from the picture below, the Missiles has interaction with Napalm. As Missiles increase, the number of enemy casualties decrease. That’s probably because the enemy move away after the first several missiles are used. As Napalm increase, the number of enemy causalities drops more quickly as Missiles increase. That’s probably because as we give more gallon Napalm to the front attack unit, the power of Missiles we use is less powerful.



As Napalm increase, the number of enemy casualties decrease. That’s probably because the enemy move away after they find we use napalms. As Missiles increase, the number of enemy causalities drops more quickly as Napalm increase. That’s probably because as we give more gallon Napalm to the front attack unit at the same time when we send Missiles, the enemies are more likely to move away soon.



As media increase per 1%, the number of enemy casualties decrease by 2.107412e+02%. That’s probably because as more percentage of suppression placed on the media coverage of the war, it also lowers troop morale.



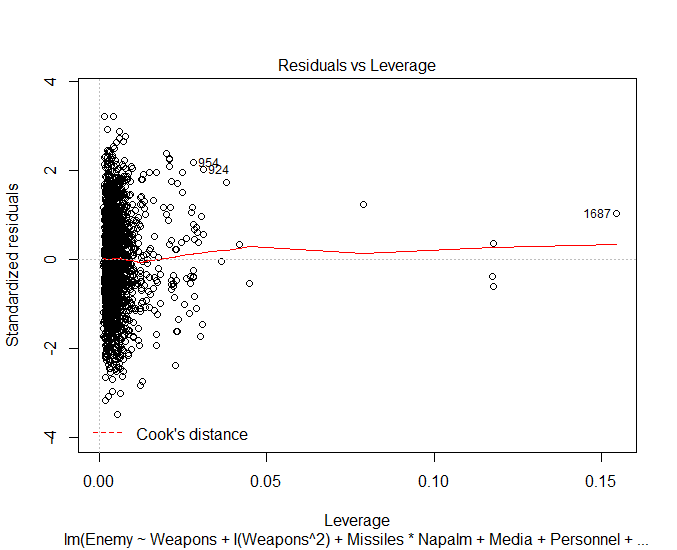
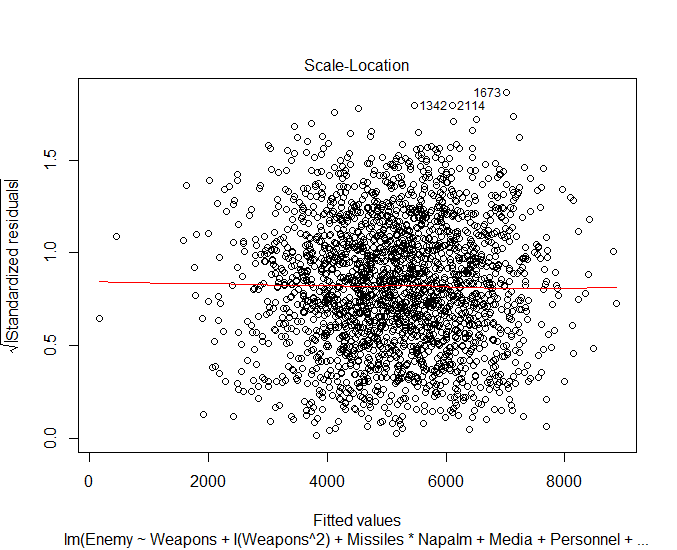
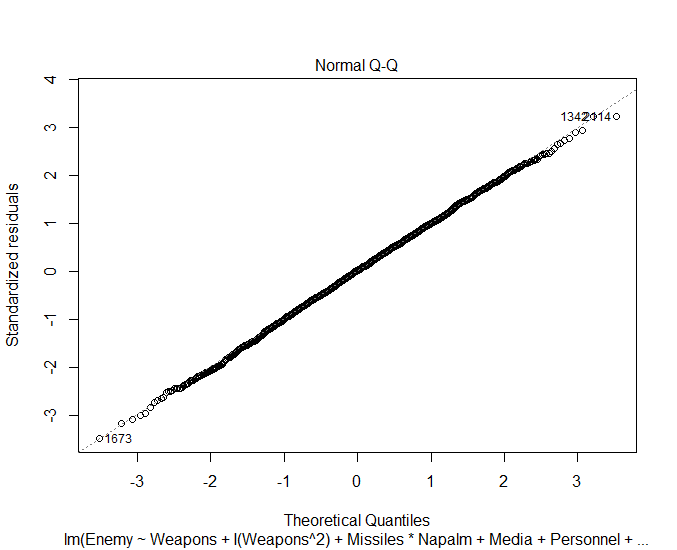
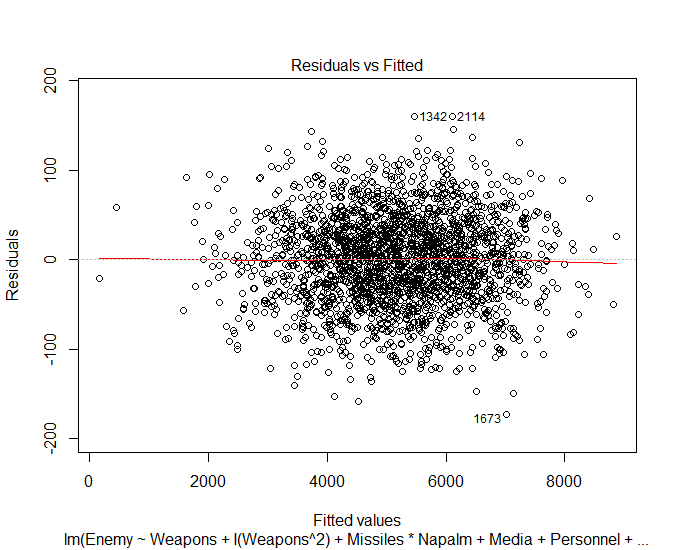
From the picture above we can see that, the effect of Personnel is cubic. As Personnel increase, the number of enemy casualties increase at first, decrease then, and increase again. That’s probably because the assist from personnel is helpful to defeat enemies in total. But as numbers of personnel increase around 28, they are likely to have disagreements inside, and this increase mistakes they made. As more personnel joining, the power of their assist being more effective.

As temperature increase per Fahrenheit, the number of enemy casualties increase by 1.457744e+00. That’s probably because the weather in Latveria is quite hot. When it’s hotter, the environment for enemy is becoming worse and more enemies die.

If we place IG88 guidance system, the number of enemy casualties will decrease by 2.484868e+02. From the result, there are probably something wrong with the IG88 guidance system that it makes wrong decisions.

In my model, the effects of Stock, Terrorism, Firepower, Payload, Bomb, First Aid and Spies on the number of civilian casualties are not significant according to the p-values. So, we don’t have enough evidence to say that Stock, Terrorism, Firepower, Payload, Bomb, First Aid and Spies have effects on the number of civilian casualties.

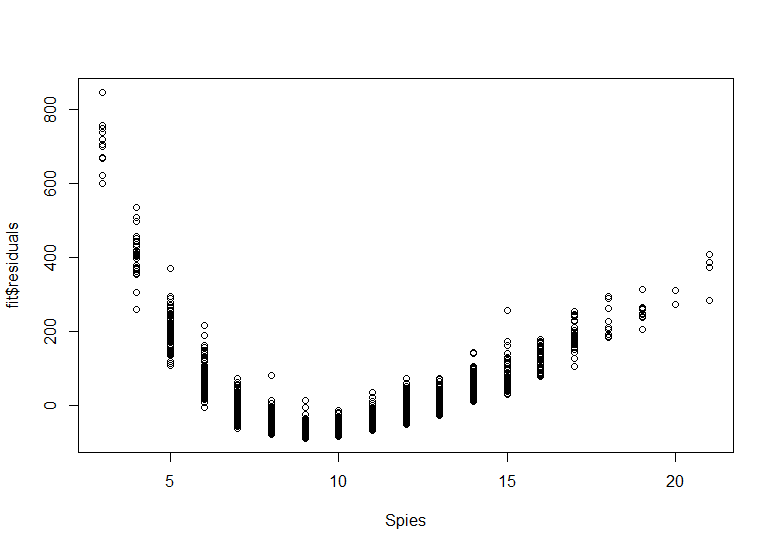
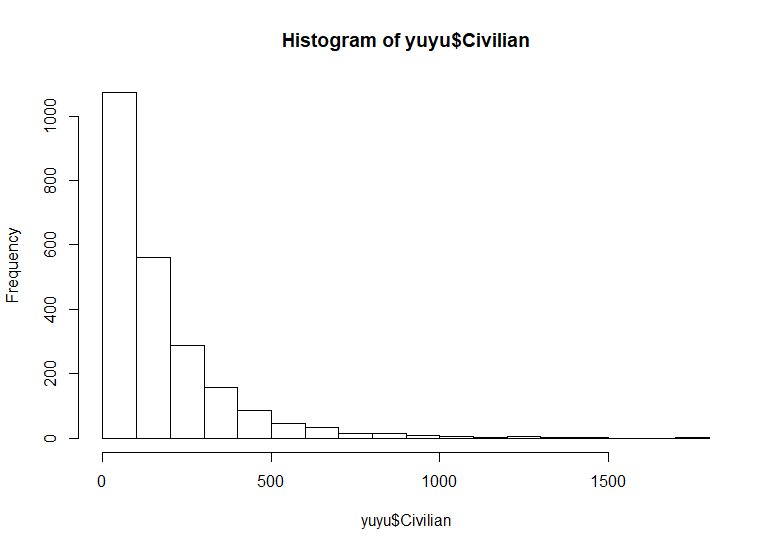
To maximize the number of enemy casualties, we have to minimize the number of missiles, Napalm, media, maximize temperature, either minimize or maximize weapon and personnel but donlt keep them at middle level, and don’t installed IG88. Which means, we should use as less Missiles and Napalm as possible, and reduce the percentage of suppression placed on the media coverage of the war. At the same time, choose to attack enemies when the temperature is highest. And if possible maximize the budget for weapons or minimize it. Also, choose 28 or 34 personnel to assist the soldiers from their desks in Washington D.C. Last but not least, don’t install IG88.



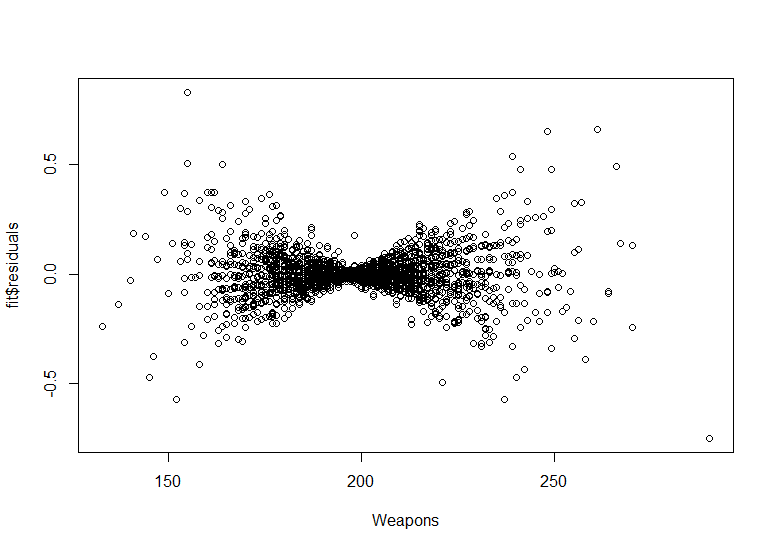
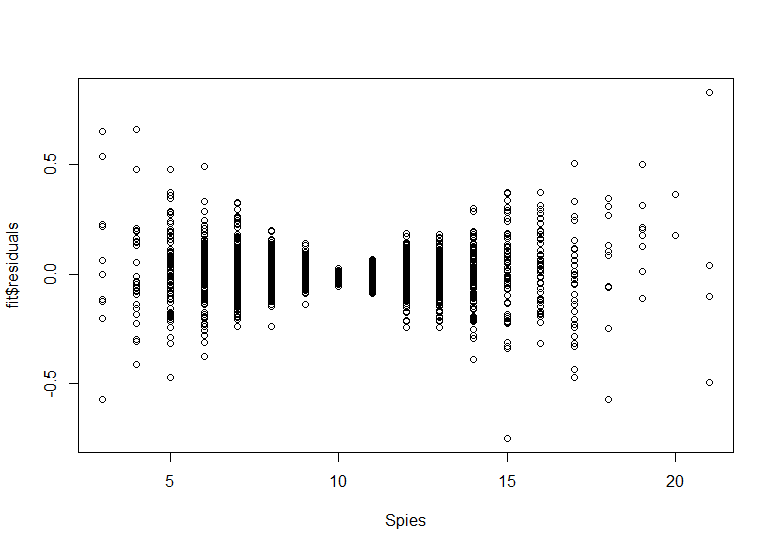
I know I have a good model because the residuals looks randomly and all very nice. All the P-values are significant at an alpha of 0.05. Besides, the prediction is pretty accurate as I got a standard error as small as 49.86 on 2285 degrees of freedom. Which means the prediction of my model will possibly have a difference under 49.86 with the real data, and this is allowed by the model. I use the first row of data to make a prediction of enemy casualties with a result of 3980.65, which is only 49.77 from the actual enemy casualties of 3923. And this smaller than the standard error given by the analysis that it’s foreseeable that we can accept it. What’s more, my R2 is as big as 0.9983 which means my prediction equation has accounted 99.83% of variability in the data. The confidence for smallest beta was [9.825295e+00, 1.102421e+01], which means we have 95% confidence to say that the effect of Personnel^3 on the number of Enemies Casualties will be within [9.825295e+00, 1.102421e+01].

## How to minimize the number of civilian casualties?

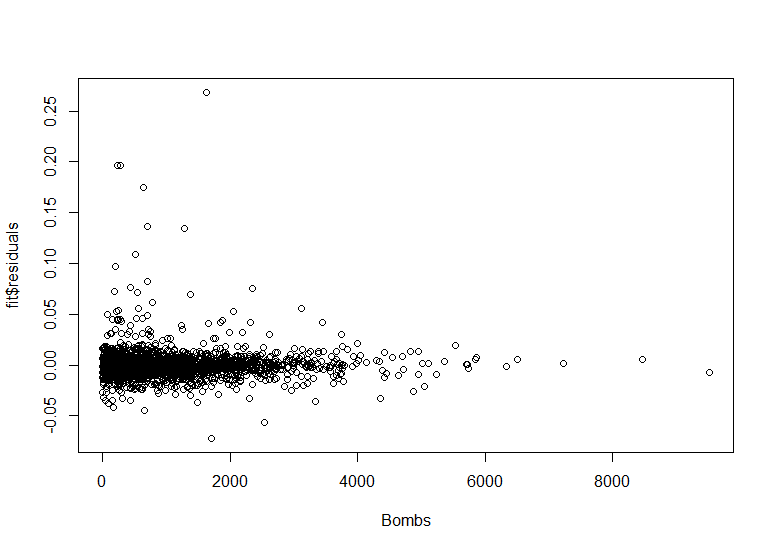
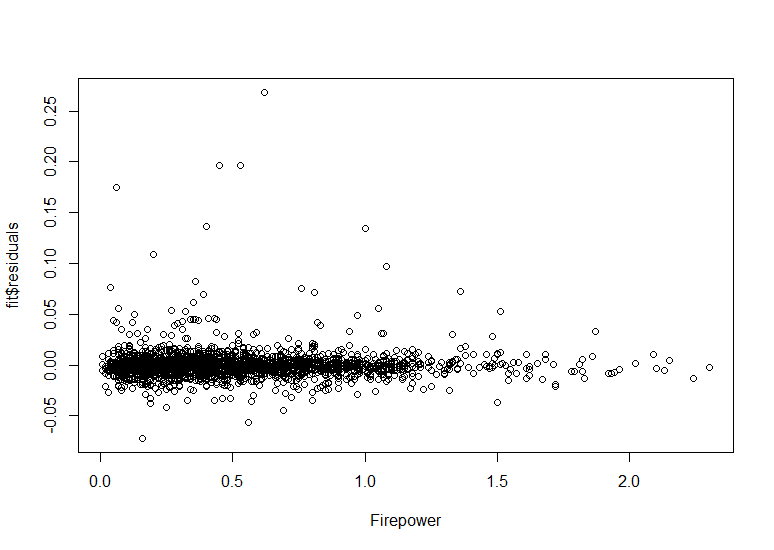
In the second part, I hope to develop a model to predict the number of civilian casualties based on other factors. I used R-studio to analyze the data at an alpha of 0.05. When I look at the data, I find most time there are 0-500 civilian casualties. There are only three plots that the number of civilian casualties are bigger than 1500, to better analyze I take them as mistakes and delete them. I made a linear model with all the variables to predict the enemy causalities, but the residual plots all looks bad so I know I made a wrong model. To fix the model, I looked at the residual plots against each variable. And I find a clear pattern of the residuals against Spies, which suggested that there is a logged effect of civilian casualties’ numbers.

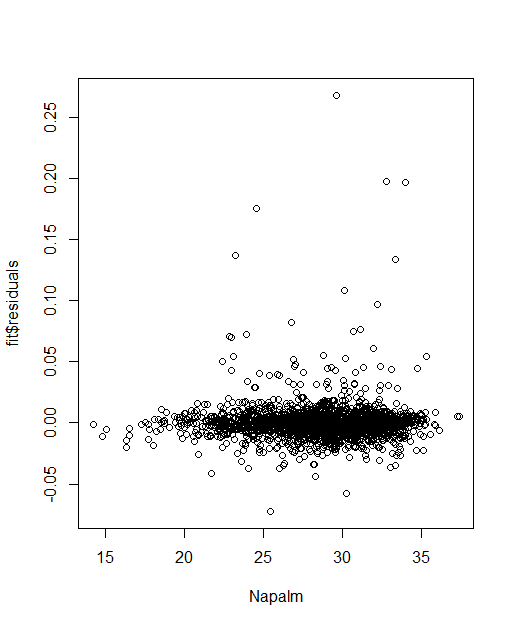


I add the logged effect of civilian casualties and looked at the residuals against each variables. I find the residual plots against both Weapons and Spies looks like a bowtie shape, so I realized that they are interacting with each other. I added the numerical interaction between them and looked at the residual plots again.

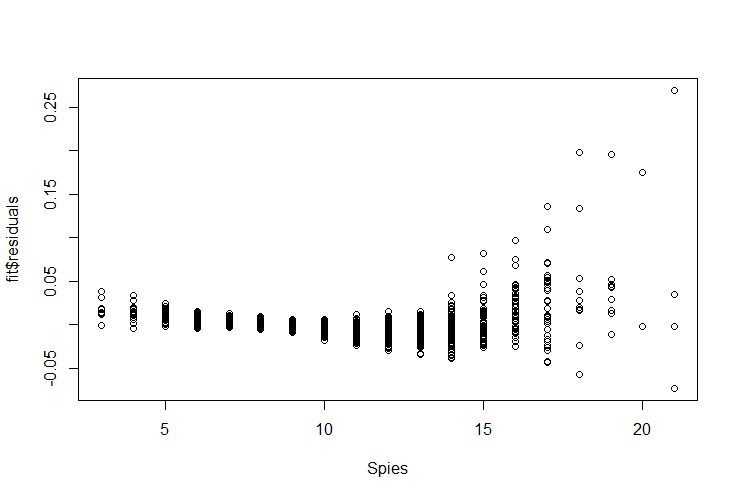


Then I find that the residual plots against Firepower and Bombs all shows a pattern of periodical wave, which suggests that there are effect of sin. I know I am correct because residuals against them looks better after I add the effect of sin to them.



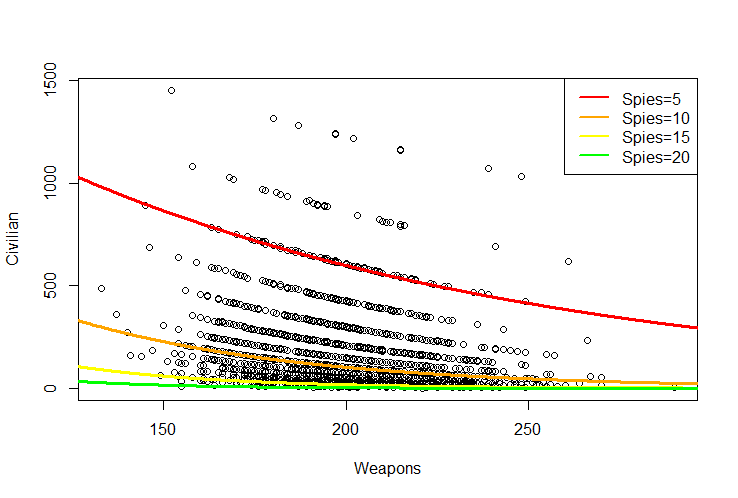


Furthermore, I find there is a pattern of cubic in the variables against Napalm although it’s very subtle. So, I add the cubic effect to it. However, the residual plot of Spies still looks weird, so I try several interactions of spies with other variables. And surprisingly I find that Sin(Bombs), personnel, Missiles, and payload are all possibly interacting with Spies according to the p-values which are significant. So, I add these four interactions into my model.

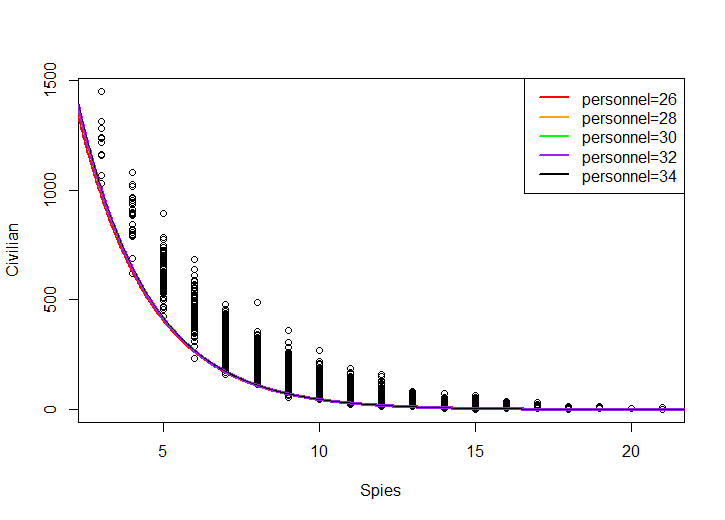
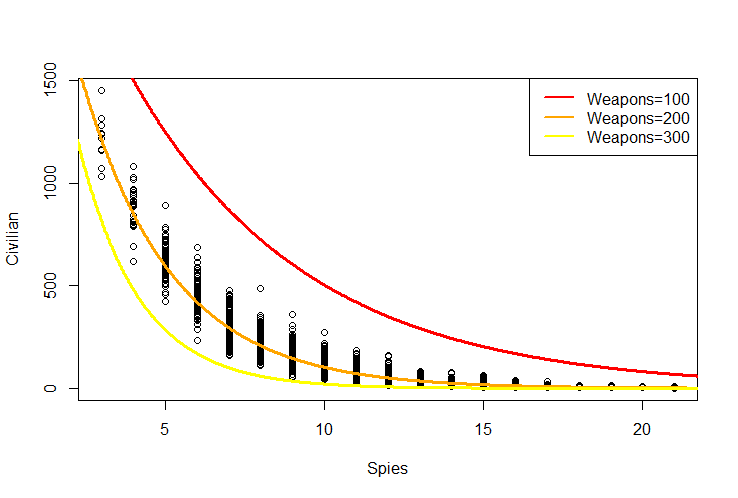


My final equation is as following:

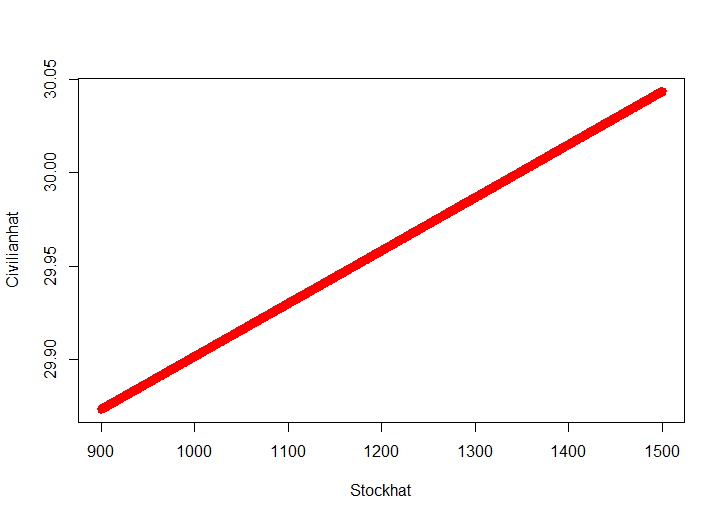
Civilian<-exp(7.554502e+00+1.156710e-03\*Weapons+8.036488e-03\* Spies+9.308792e-06\*Stock+3.514369e-06\*Payload + 1.108670e-02\*TerrorismLow+2.180982e-02\*TerrorismMedium+3.068151e-03\*sin(Firepower)-3.029101e-03\*sin(Bombs)+2.350628e-06\*Bombs+5.780145e-03\*Personnel-5.110269e-04\*Missiles+2.099480e-02\*Napalm-8.195829e-04\*I(Napalm2)+1.035063e-05\*I(Napalm3)-1.705624e-03\*Weapons\*Spies-3.876748e-07\*Spies\*Payload+3.333599e-04\*Spies\*sin(Bombs)-2.842763e-07\*Spies\*Bombs-6.085789e-04\*Spies\*Personnel+5.421196e-05\*Spies\*Missiles)



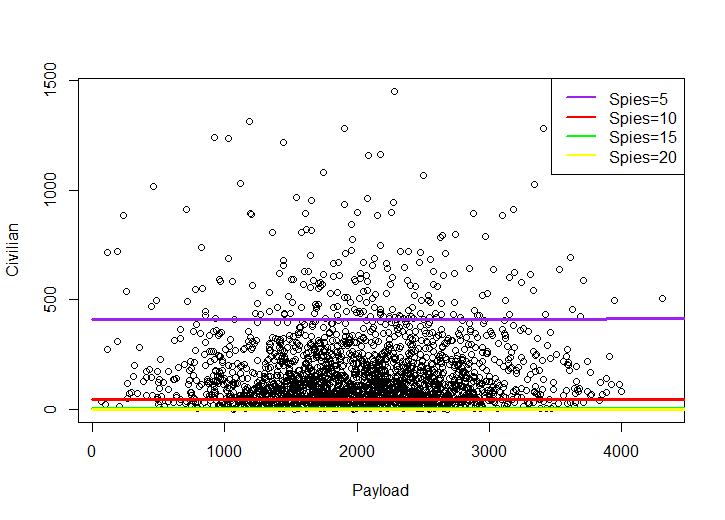
As we can see from the picture, as Weapons increase, the number of civilian casualties decrease. And it decrease fast at first, and then slow down. That’s probably because as more budget are being put for weapons, the more likely that the civilians are going to be protect safely. Furthermore, the effect of weapons is interacting with Spies. As the number of Spies increase, the effect of weapons is becoming less powerful, and decrease more slowly. That’s probably because, with more spies we are more likely to know the enemy’s battle strategy that we can protect our civilian better, and at this time, the protect from Weapons is not as significant as before.



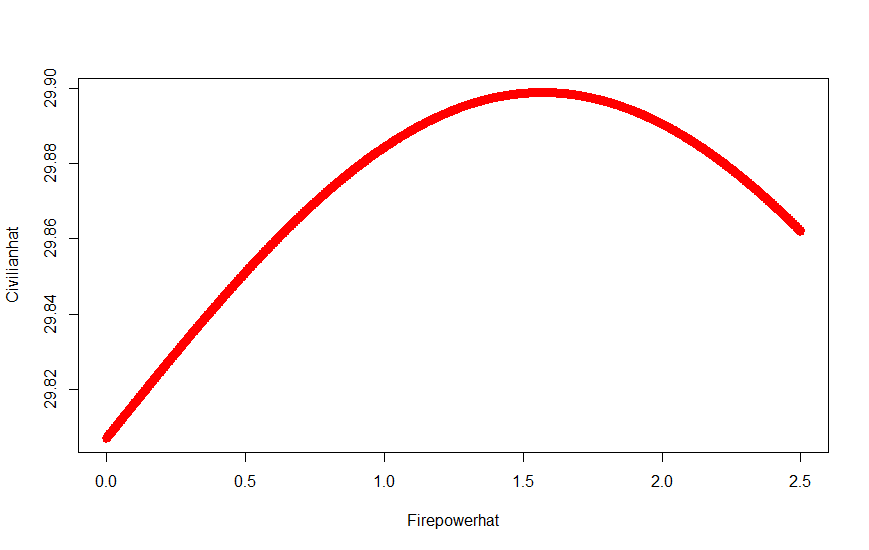
Look at the picture we may find that, as Spies increase, the number of civilian casualties decrease. And it decrease fast at first, and then slow down. That’s probably because as more spies are being sent, we are more likely to know the enemy’s battle strategy that we can protect our civilian better. Furthermore, the effect of weapons is interacting with Spies. As the number of Weapons increase, the effect of Spies is becoming more powerful, and decrease more quickly. That’s probably because, with more weapons we are able to make use of Spies better and protect our civilians better. At the same time, the Spies is also interacting with Sin(Bombs), personnel, Missiles, and payload. But as shown in the picture, the effect of other variables on Spies are so tiny that we can hardly tell their difference.



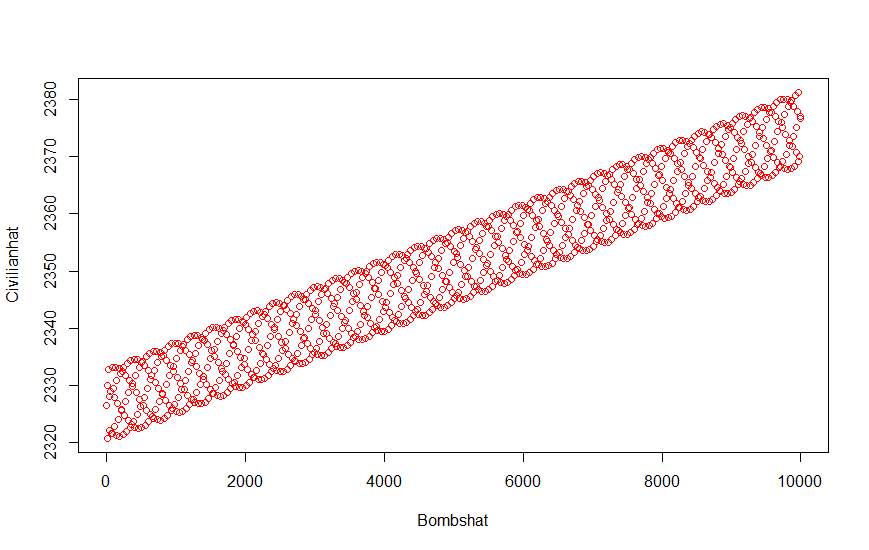
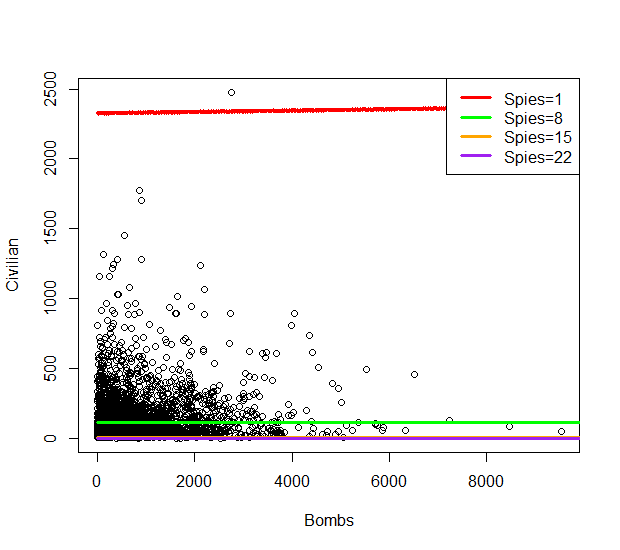
Stock also has an influence on the number of civilian casualties. As stock increase by 1 unit, the number of civilian casualties increase by |1-e9.308792e-06 |%. That’s probably because the stock market plays a role in the morale of the troop. When the stock value is high, less warriors want to works hard to protect our civilians.



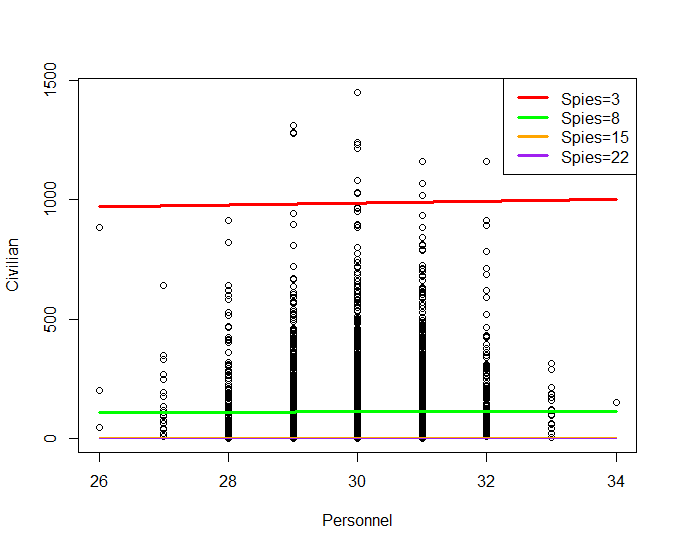
The effect of Payload is interacting with Spies. When spies are low like 5, as Payload increase, the number of civilian casualties increase. But when spies are higher, as Payload increase, the number of civilian casualties decrease. That’s probably because, when there are not too much spies, payload are more likely to hurt the civilians based on the wrong information. But when there are more spies, the payload plays an important role in attacking enemies and protect civilians. But overall, the effect of payload is not obvious.



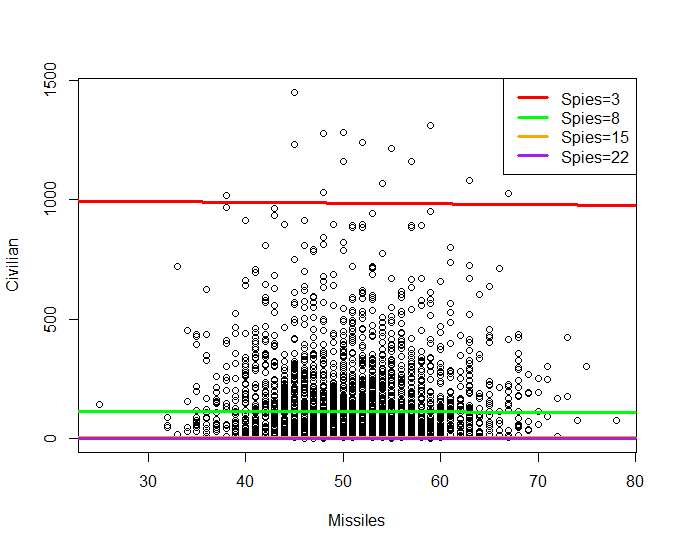
The effect of Firepower is quadratic. When the Firepower increase, the number of civilian casualties increase at first, and decrease then. That’s probably because as the percentage of weapons towards enemies are lower than 1.5, it’s likely to hurt the civilians accidentally. Only when it’s higher than 1.5, it will better attack the enemies and protect the civilians.



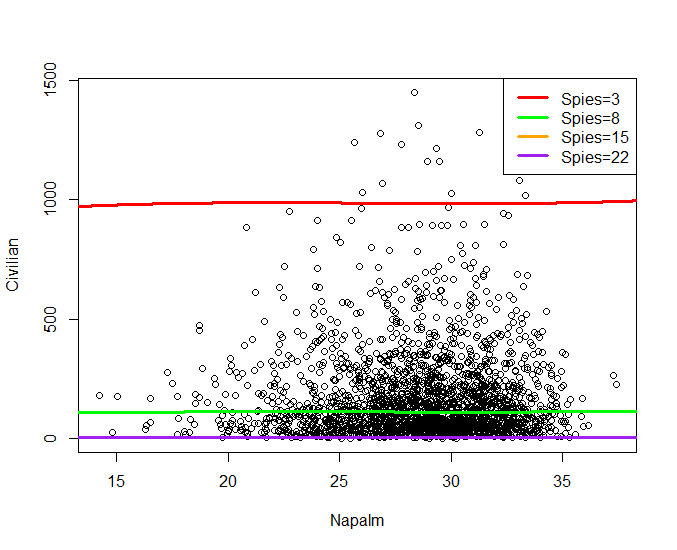
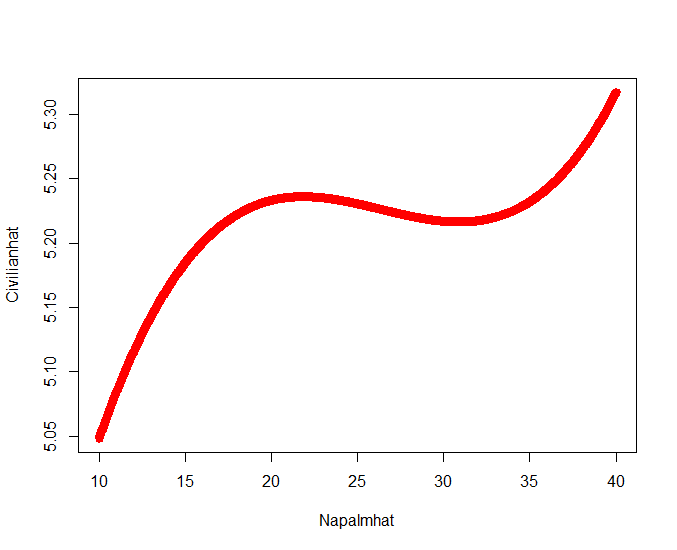
The effect of Bombs is interacting with Spies. When the spies number are low like 8, as the number of Bombs increase, the number of civilian casualties increase. As the spies number are high like 15, as the number of Bombs increase, the number of civilian casualties decrease. Also, the effect of Bomb is sined. We can see as Bombs changes, the number of civilian casualty’s changes in chained circles. That’s probably because the effect of Bombs are hard to predict and varies a lot that it gives a range of the values for bombs.



Like Bombs, the effect of Personnel is also interacting with Spies. When the spies number are low like 8, as the number of Personnel increase, the number of civilian casualties increase. As the spies number are high like 15, as the number of Personnel increase, the number of civilian casualties decrease. It’s possible that the Spies has such a strong effect, when there are a lot spies, the personnel’s effect is not very strong. When there are more spies, the personnel are more likely to make right decision based on the information provided by Spies.



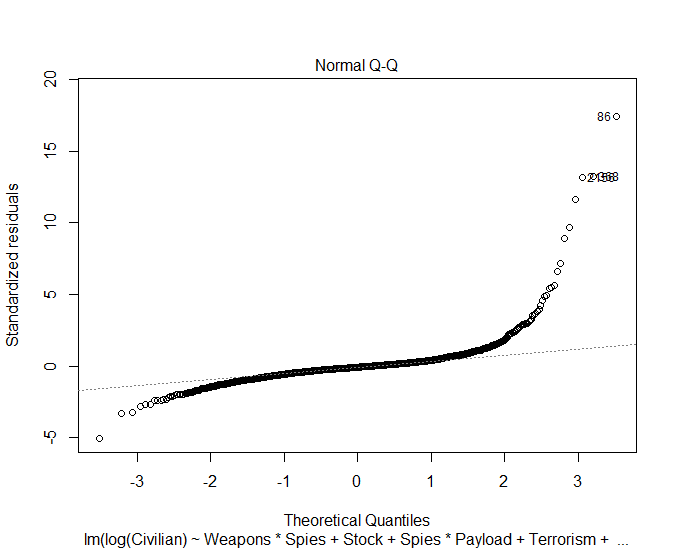
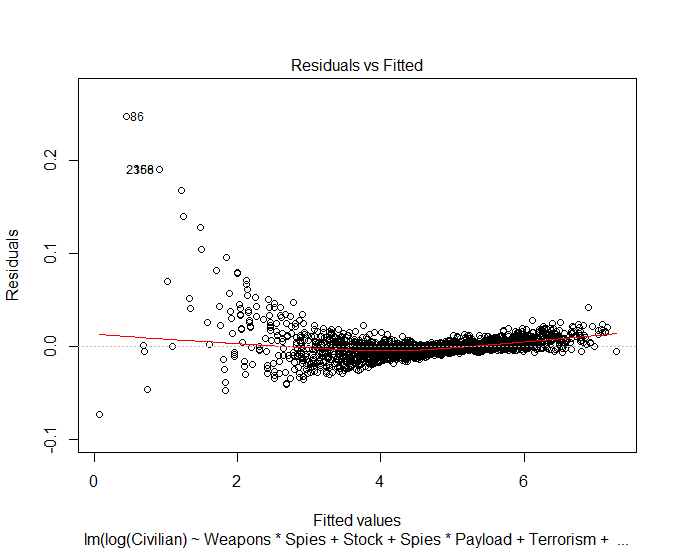
Similarly, the effect of Missiles is also interacting with Spies. When the Spies number are low like 8, as the number of Missiles increase, the number of civilian casualties decrease. As the spies number are high like 15, as the number of Missiles increase, the number of civilian casualties increase. As we can see from the picture, when the number of Spies is low, there is a high amount of civilian casualties. As the spies number increase to 8, the amount of civilian casualties decrease in a large degree. This suggests that Spies is crucial to decrease the amount of civilian casualties.

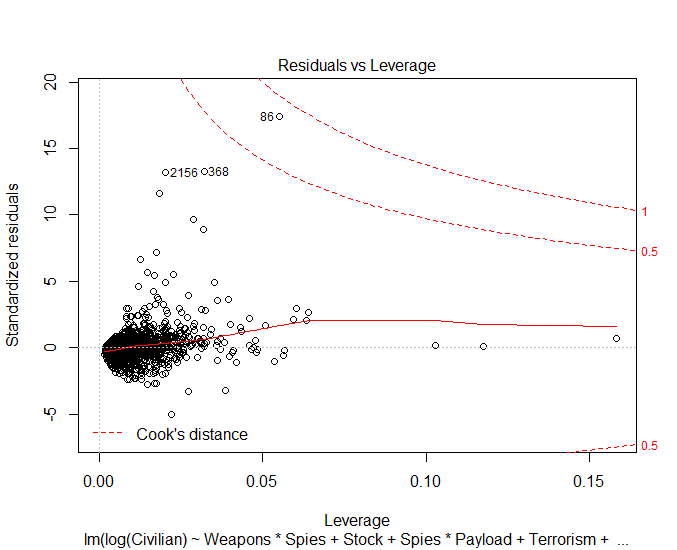
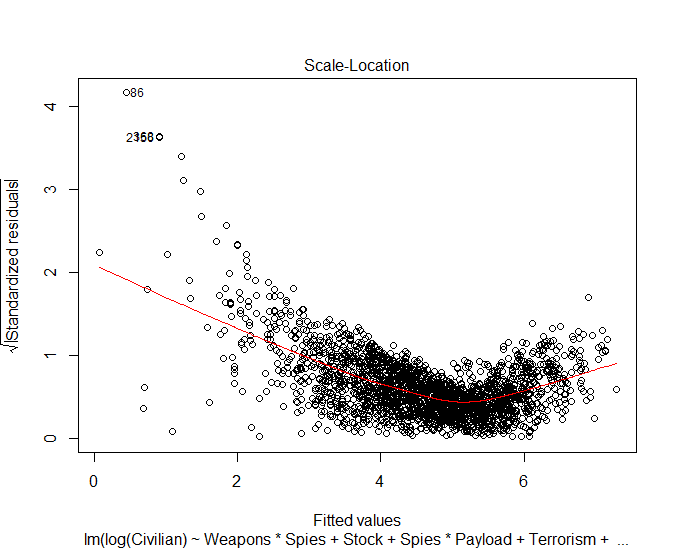


The effect of Napalm is cubic. As the number of Napalm increase, the number of civilian casualties show a trend of increasing in total. But it shows a slight trend of decrease when the number of Napalm are around 20-35, probably at this time, the Napalm has a strong effect in hurting the enemies and protect civilians. Similarly, the Napalm is interacting with Spies. Although it shows a trend of increasing no matter how many Spies we have, the increasing trend slow down when there are more Spies. That’s probably the information from Spies protect the civilian better and thus protect them from getting hurt by Napalm.

In my model, the effects of First Aid, Temperature, Media and the use of IG88 on the number of civilian casualties are not significant according to the p-values. So, we don’t have enough evidence to say that First Aid, Temperature, Media and the use of IG88 have effects on the number of civilian casualties.

To minimize the number of civilian casualties, the most crucial is to increase the number of Spies and keep it at a higher number than 8, which significantly decrease civilian casualties. At this time, maximize budget on weapons, the number of Personnel, Payload, Bombs, minimize the use of Napalms the value of Stocks, the use of Payloads, and the use of Missiles.





I know I have a good model because the residuals looks much better than they originally were. I know that residuals still looks wired, but I have made the best to make it looks better. Also, I think all the other residuals are being folded brought by the interaction between Weapons and Spies. All the P-values are significant at an alpha of 0.05. Besides, the prediction is pretty accurate as I got a standard error as small as 0.01464 on 2274 degrees of freedom. Which means the prediction of my model will have a difference under 0.01464 compared to the real data. I use the first row of data to make a prediction of enemy casualties with a result of 66.55158, which is only 0.006715333 from the actual enemy casualties of 67 after being logged. And this smaller than the standard error given by the analysis that it’s foreseeable that we can accept it. What’s more, my R2 is as big as 0.9998 which means my prediction equation has accounted 99.98% of variability in the data. The confidence for smallest beta was [-1.322294e-03, -3.148770e-04], which means we have 95% confidence to say that the effect of Napalm^3 on the number of civilian causalities will be in [-1.322294e-03, -3.148770e-04].

So, with my report, we may find that there are several factors which influence the number of enemy casualties, and based on my prediction, we may able to control those factors and thus maximize the number of enemy casualties and minimize the number of civilian casualties. And through this we can develop the best battle strategy and win the war. Besides, my study point out that there may be something wrong with IG88 guidance system that it makes wrong decision, so further study can works on this to figure out the default. Also, my study can only minimize the number of civilian casualties and maximize the number of enemy casualties separately, the strategy based on this may not able to fulfill both at the same time. Further study can combine this two in one model and thus develop the best strategy to win the war.

Appendix

Enemy

summary(fit)

Call:

lm(formula = Enemy ~ Weapons + I(Weapons^2) + Missiles \* Napalm +

Media + Personnel + I(Personnel^2) + I(Personnel^3) + Temperature +

IG88, data = yuyu)

Residuals:

Min 1Q Median 3Q Max

-172.748 -32.870 1.049 33.787 160.243

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -2.650e+05 8.168e+03 -32.440 < 2e-16 \*\*\*

Weapons -1.966e+01 7.049e-01 -27.896 < 2e-16 \*\*\*

I(Weapons^2) 5.736e-02 1.744e-03 32.882 < 2e-16 \*\*\*

Missiles 4.268e+00 1.302e+00 3.278 0.00106 \*\*

Napalm 4.517e+00 2.319e+00 1.948 0.05157 .

Media -2.106e+02 8.624e+00 -24.425 < 2e-16 \*\*\*

Personnel 2.803e+04 8.202e+02 34.175 < 2e-16 \*\*\*

I(Personnel^2) -9.379e+02 2.744e+01 -34.177 < 2e-16 \*\*\*

I(Personnel^3) 1.042e+01 3.058e-01 34.084 < 2e-16 \*\*\*

Temperature 1.457e+00 6.273e-02 23.233 < 2e-16 \*\*\*

IG88Installed -2.485e+02 2.085e+00 -119.179 < 2e-16 \*\*\*

Missiles:Napalm -4.594e+00 4.518e-02 -101.693 < 2e-16 \*\*\*

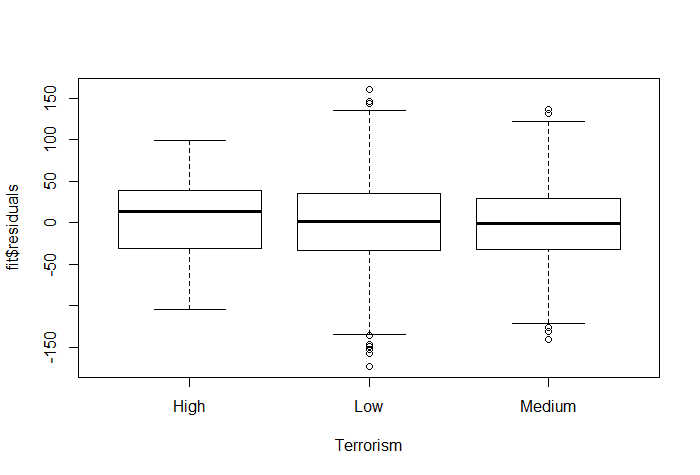
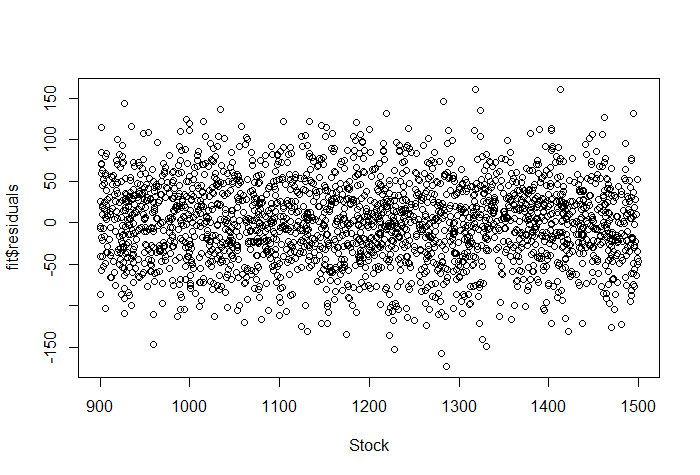
---

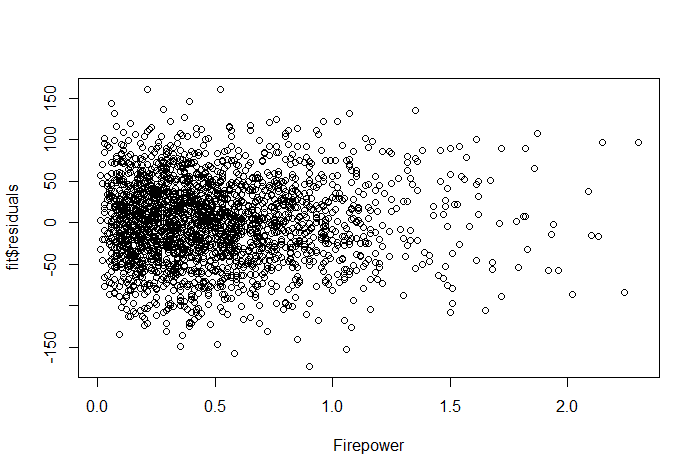
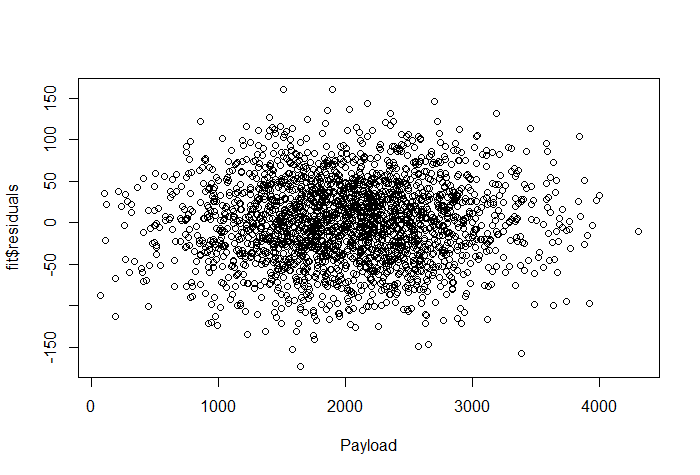
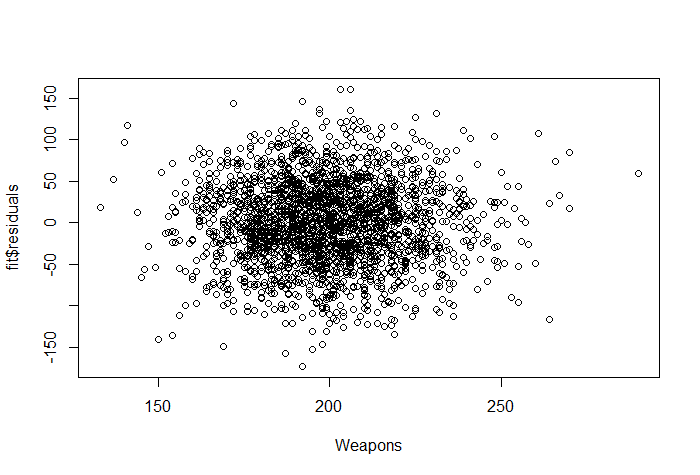
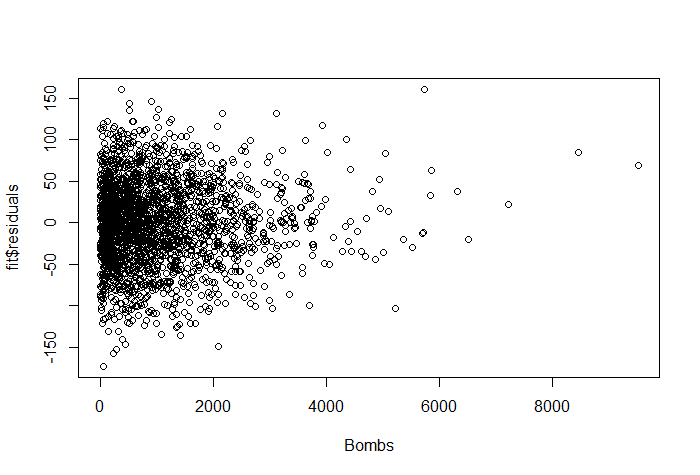
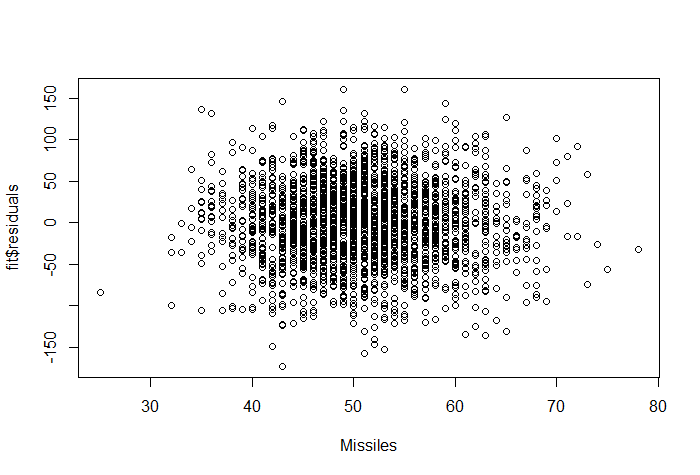
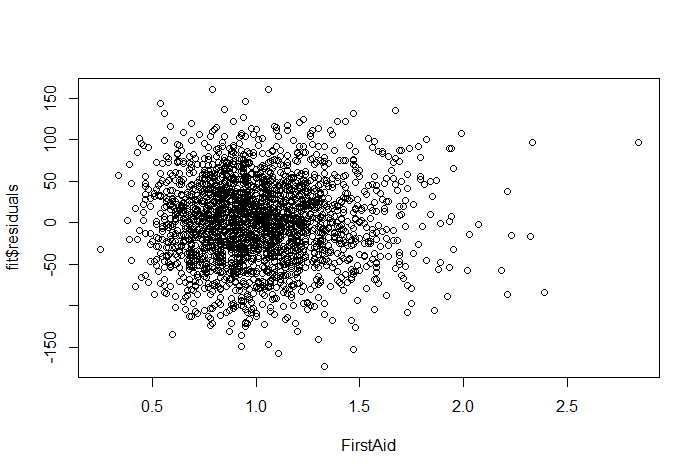
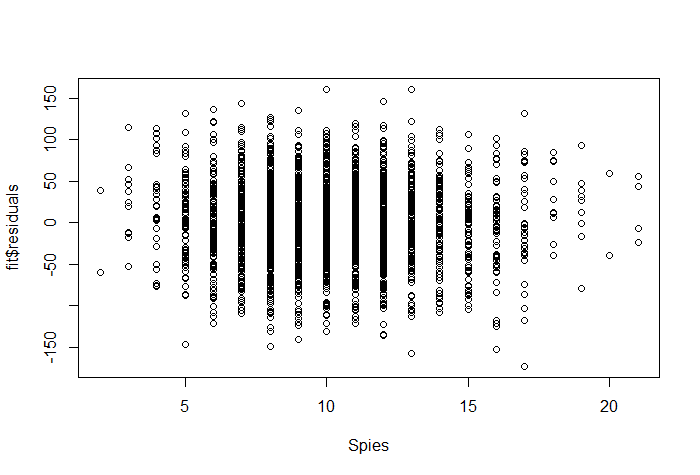
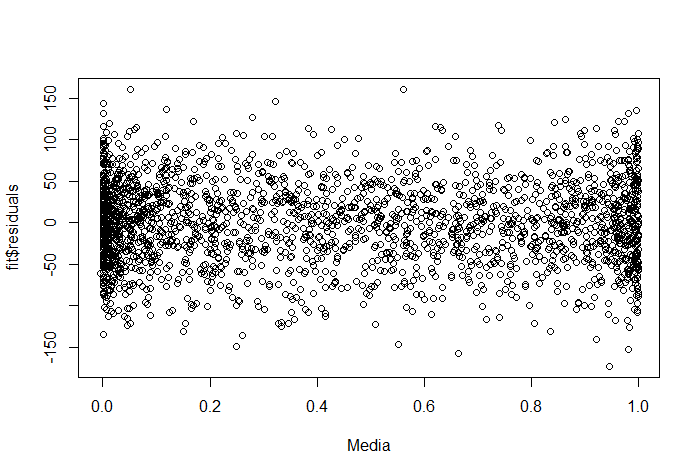
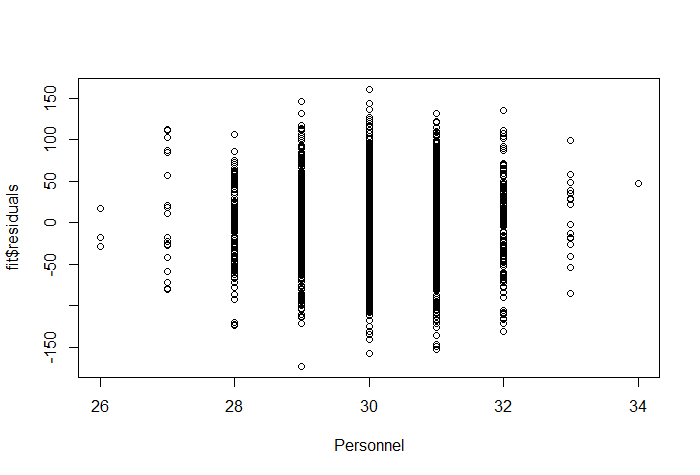
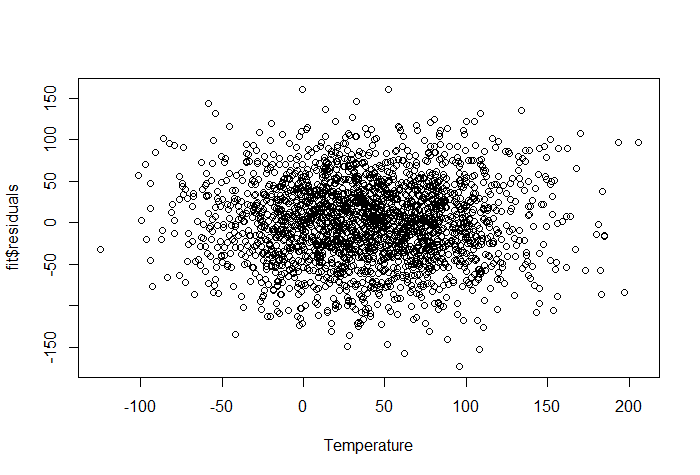
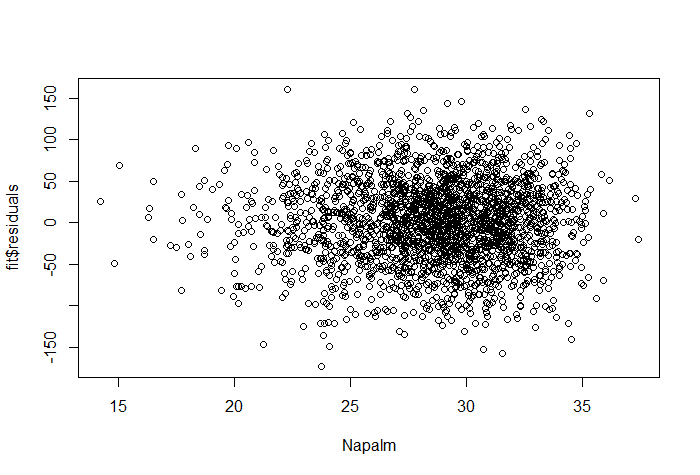
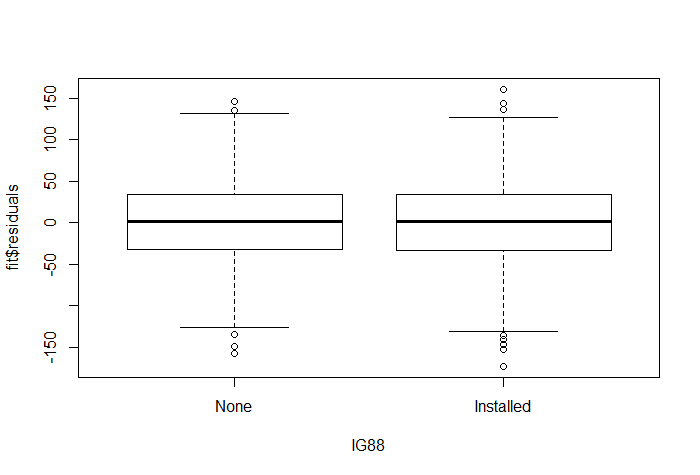
Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 49.86 on 2285 degrees of freedom

Multiple R-squared: 0.9983, Adjusted R-squared: 0.9983

F-statistic: 1.196e+05 on 11 and 2285 DF, p-value: < 2.2e-16



summary(fit2)

Call:

lm(formula = log(Civilian) ~ Weapons \* Spies + Stock + Spies \*

Payload + Terrorism + sin(Firepower) + Spies \* sin(Bombs) +

Spies \* Bombs + Spies \* Personnel + Spies \* Missiles + Napalm +

I(Napalm^2) + I(Napalm^3), data = yuyu)

Residuals:

Min 1Q Median 3Q Max

-0.072693 -0.005234 -0.001055 0.003152 0.247839

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 7.553e+00 6.928e-02 109.028 < 2e-16 \*\*\*

Weapons 1.151e-03 5.487e-05 20.974 < 2e-16 \*\*\*

Spies 8.415e-03 3.205e-03 2.625 0.008714 \*\*

Stock 9.477e-06 1.779e-06 5.326 1.10e-07 \*\*\*

Payload 4.180e-06 1.575e-06 2.653 0.008029 \*\*

TerrorismLow 1.102e-02 1.686e-03 6.538 7.65e-11 \*\*\*

TerrorismMedium 2.178e-02 1.732e-03 12.575 < 2e-16 \*\*\*

sin(Firepower) 3.011e-03 1.194e-03 2.522 0.011750 \*

sin(Bombs) -2.837e-03 1.542e-03 -1.840 0.065868 .

Bombs 2.284e-06 1.108e-06 2.060 0.039491 \*

Personnel 5.872e-03 1.021e-03 5.753 9.97e-09 \*\*\*

Missiles -4.877e-04 1.545e-04 -3.156 0.001620 \*\*

Napalm 2.063e-02 6.790e-03 3.039 0.002404 \*\*

I(Napalm^2) -8.055e-04 2.564e-04 -3.141 0.001703 \*\*

I(Napalm^3) 1.018e-05 3.189e-06 3.191 0.001437 \*\*

Weapons:Spies -1.705e-03 5.162e-06 -330.319 < 2e-16 \*\*\*

Spies:Payload -4.477e-07 1.519e-07 -2.947 0.003246 \*\*

Spies:sin(Bombs) 3.154e-04 1.472e-04 2.143 0.032231 \*

Spies:Bombs -2.787e-07 1.041e-07 -2.676 0.007510 \*\*

Spies:Personnel -6.164e-04 9.735e-05 -6.332 2.91e-10 \*\*\*

Spies:Missiles 5.206e-05 1.483e-05 3.510 0.000458 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.01464 on 2274 degrees of freedom

Multiple R-squared: 0.9998, Adjusted R-squared: 0.9998

F-statistic: 6.235e+05 on 20 and 2274 DF, p-value: < 2.2e-16

