Eric He

Regression and Multivariate Data Analysis

Homework 6

Overview of the Topic

World of Tanks is a computer game which pits two teams of 15 players against each other with the win conditions of either destroying the entire enemy team or capturing the base. Players get to drive over a hundred different tanks from nine different nations.

Tanks can be further divided by type and tier. The five tank types are light, medium, heavy, tank destroyer, and artillery. Each tank type has its own strengths and weaknesses which given them specialized roles in the game. Light tanks are scouts which spot enemy tanks for their team to snipe. Heavy tanks are slow and unwieldy, but heavily armored. Medium tanks occupy a space in the middle which allows them to hunt down light tanks and outmaneuver (but not outgun) other tank types. Tank destroyers tend to have no turrets, allowing them to pack greater firepower at the expense of being helpless when flanked; as such, they tend to stay back and snipe. Artillery provides long range firepower with the ability to fire shells in a curved arc that allows them to (sometimes) bypass the cover a tank is hiding behind by having the shell drop down on top of the enemy.

There are ten tiers of tanks, with higher tier tanks tending to have more health, better damage, and thicker or more sophisticated armor. Higher tier tanks correspond to later generations of tanks. The tier 1 Leichttraktor, for example, is a German tank produced from 1930 to 1934 while the tier 10 M48A1 Patton is an American tank first developed in 1950.[1] Including modern tanks is not viable for the game because some nations, such as Japan and Germany, stopped developing tanks after World War 2, and because of a paradigm shift in military science

[1] http://wiki.wargaming.net/en/World_of_Tanks

and arms technology which phased out the need for both light and heavy tanks. Many tanks in the game are only experimental ones that never made it past the design stage in real life[2].

My data analysis will focus on 92 games where I played with the Type 59, a tier 8 Chinese medium tank manufactured from 1958 to 1987. Type 59s are the tanks seen in the famous Tiananmen Square tank man video, a picture of which is shown below.[3]

The Type 59 could only be bought for real money and has the benefits of getting bonus in-game monetary rewards and preferential matchmaking; while other tier 8 tanks may face tier 10 tanks every once in a while, the Type 59 can only be matched against tier 9 tanks at the maximum.

Though the Type 59's armor is not thick enough to block most shots head on, the armor (this can be seen in the above picture) is angled or "sloped." Sloped armor means that rather than

[2] An example being the Batignolles-Châtillon 25t, a French medium tank of which only two prototypes were made.
[3] Source: http://lens.blogs.nytimes.com/2009/06/03/behind-the-scenes-tank-man-of-tiananmen/

having a shell punch through the armor, there is a possibility that an enemy shot can be deflected away because its impact trajectory is not orthogonal to the armor plating. The sloped armor gives the Type 59 the ability to occasionally "bounce" shots from tanks with extremely high firepower, which infuriated so many players that the developers removed the Type 59 from the in-game store in 2012. Due to these features, especially the last one, I play the Type 59 the most of all my tanks. The combination of sloped armor, maneuverability, and an inaccurate but rapid firing gun gives the Type 59 a close range, brawling play style.

Overview of the Data

Descriptive Statistics: W/L, Kills, Damage, Assisted, Credits

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
W/L	92	0	0.5761	0.0518	0.4969	0.0000	0.0000	1.0000	1.0000	1.0000
Kills	92	0	1.217	0.133	1.274	0.000	0.000	1.000	2.000	5.000
Damage	92	0	1397.7	82.1	787.6	0.0	754.8	1403.5	1985.3	3426.0
Assisted	92	0	512.2	52.4	502.8	0.0	119.8	356.5	749.8	2521.0
Credits	92	0	38558	1538	14754	9292	27849	38929	49499	76307

	Tank	W/L	Kills	Damage	Assisted	Survived	Credits
1	Type 59	1	3	2792	0	1	55578
2	Type 59	1	4	2027	189	1	48796
3	Type 59	1	0	723	447	0	30573
4	Type 59	1	1	2774	610	0	64261

There are 92 data points, each corresponding to a different game in the Type 59. Data recorded for each game include the amount of kills, damage, assisted damage, and credits I got per game, and if my team won or lost the game. Though draws where neither team wins are possible, they are fairly uncommon and none are in this dataset.

The W/L variable represents whether I won or lost the game, with a 1 corresponding to a win and a 0 corresponding to a loss. The data shows of the 92 data points, I won 57.61%, or 53 of them. This win ratio is slightly inflated because I occasionally log off after having my tank destroyed mid-game without recording the data, and when my tank is destroyed before the game is finished, the odds are higher that I lose the game. This is a problem in that there is less data for the losses.

The Kills variable counts how many tanks I destroyed during the game. The mean of 1.217 shows that on average, I destroy 1.217 tanks every game. In the dataset, the maximum amount of tanks I destroyed in a single game was 5. Because destroying enemy tanks eliminates the possibility that they do any more damage, I believe more kills should be associated with a higher odds of victory.

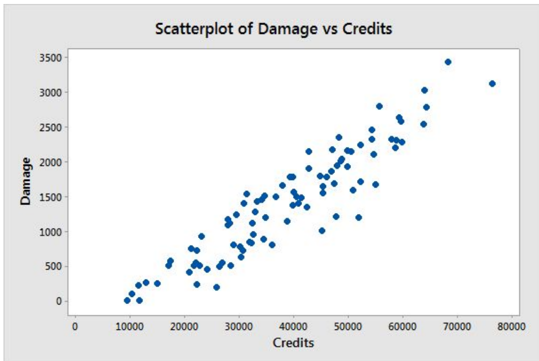
The Damage variable shows how much damage I did during the game. The average is 1397.7 hitpoints of damage dealt every game. For reference, the Type 59 has 1300 hitpoints. The same logic for kills applies to damage; higher damage should be associated with a higher odds for victory.

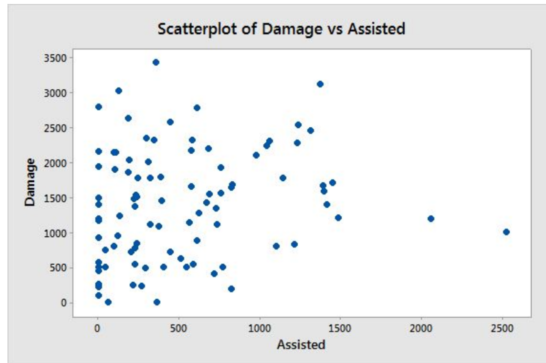
The Assisted variable represents the amount of damage my teammates did with my help. There are two ways for my teammate's damage to be counted as my assist. The first way is for my tank to be close enough to "spot" an enemy tank and reveal its location to my team. Any damage done to the enemy by my teammates which were unable to spot the enemy tank without me gets counted as assisted damage. The second type of assisted damage is damage done by my teammates when I have destroyed the tracks of an enemy tank, rendering it immobile until the enemy tank's crew fixes the tracks (in game, this takes several seconds; it would probably take longer in real life). The logic goes that if I render the enemy immobile, I give my teammates the ability to shoot weak spots or land shots that they otherwise would not have been able to make.

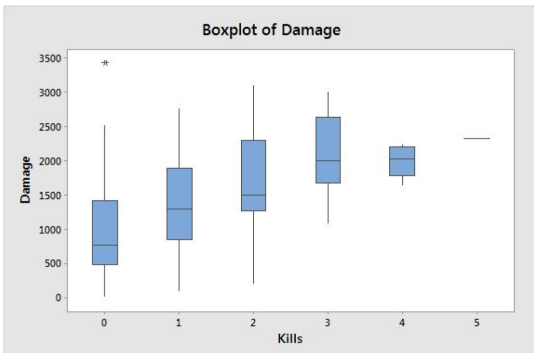
Finally, the Credits variable is the amount of in-game currency I am rewarded after each game. Credits is used to purchase or repair tanks and pay for ammunition, and the amount rewarded is based on the previously mentioned damage and assisted damage (though not kills, as I found out when doing another regression), but also on other measures I cannot record such as amount of capture points I got or the aforementioned destroying of an enemy's tracks or other modules, such as the gun, turret, fuel tanks, or engine; in other words, other relevant information that could contribute to a victory. Notably, credits awarded is not affected by the amount of kills obtained. The credits formula differs for every tank and its exact coefficients are unknown, but the components of the formula have been released[4].

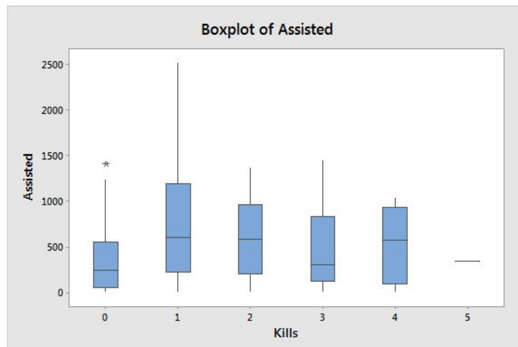
All the variables have extremely large standard deviations relative to their means, which signifies my erratic performance in this game.

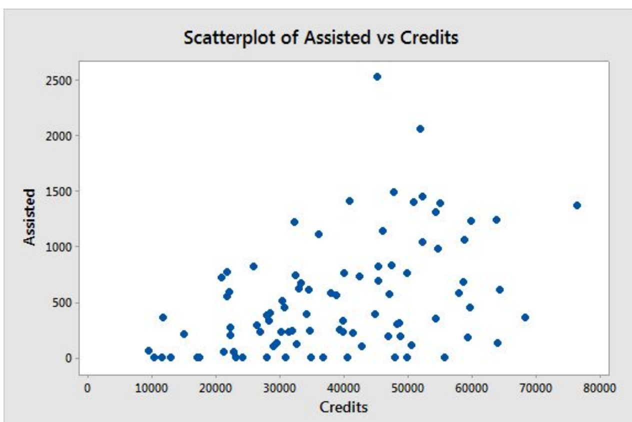
[4] However, a summary of known components can be found at http://forum.wotblitz.com/index.php?/topic/4198-how-do-they-calculate-credits-after-battles/



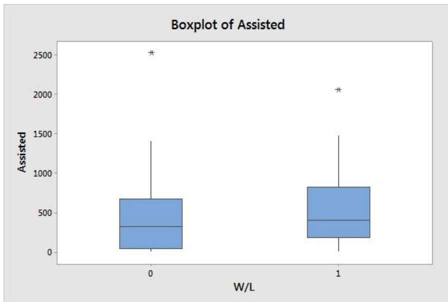


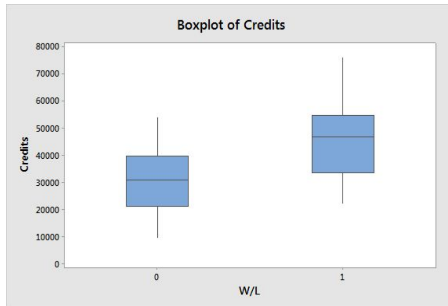


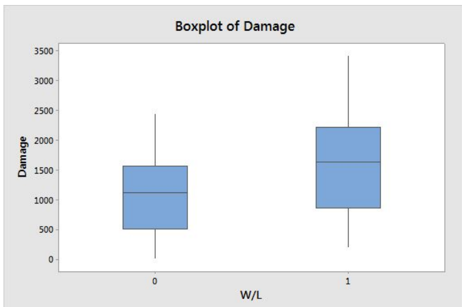




The relationships between Damage with Credits and Damage with Kills is highly correlated but apart from that there does not seem to be any multicollinearity issues. If the VIFs for either variable are too high I will experiment with different models having only one of them.







Tabulated Statistics: Kills, W/L

Rows: Kills   Columns: W/L

	0	1	All
0	21	13	34
	61.76	38.24	100.00
1	14	14	28
	50.00	50.00	100.00
2	1	12	13
	7.69	92.31	100.00
3	2	9	11
	18.18	81.82	100.00
4	1	4	5
	20.00	80.00	100.00
5	0	1	1
	0.00	100.00	100.00
All	39	53	92
	42.39	57.61	100.00

Cell Contents: Count
% of Row

All the direct relationships between the predictors appear to be clear, with a higher proportion of wins for higher values of each respective predictor. There is a high leverage point shown in the Assisted vs. W/L graph where the highest point, by far, is marked as a loss. This is data point 39, where I did 1007 damage, destroyed 1 tank, and got 2521 assisted damage. Other data not included in the dataset I am currently analyzing showed that I survived the battle, which means the enemy team was able to capture my base. As said earlier, victory through base capture is fairly rare and this battle could be said to be abnormal, but with so many data points the model should not be affected too much.

W/L ~ Kills + Damage + Assisted + Credits

Binary Logistic Regression: W/L versus Kills, Damage, Assisted, Credits

* WARNING * When the data are in the Response/Frequency format, the Residuals versus fits plot is unavailable.

Method

Link function	Logit
Residuals for diagnostics	Pearson
Rows used	92

Response Information

Variable	Value	Count	
W/L	1	53	(Event)
	0	39	
	Total	92	

Deviance Table

Source	DF	Adj Dev	Adj Mean	Chi-Square	P-Value
Regression	4	107.174	26.7935	107.17	0.000
Kills	1	2.435	2.4346	2.43	0.119
Damage	1	75.723	75.7226	75.72	0.000
Assisted	1	43.864	43.8642	43.86	0.000
Credits	1	90.643	90.6431	90.64	0.000
Error	87	18.226	0.2095		
Total	91	125.400			

The model was run using the binary response/frequency format because each row corresponded to a single game where there either was one event only (a win) or wasn't (a loss). The variables are all highly statistically significant except for the Kills variable, which has a relatively low chi-squared value of 2.43, which corresponds to a p-value of 0.119. The chi-squared value for the model as a whole is 107.17, which means the model is extremely statistically significant.

```
Model Summary

Deviance   Deviance
    R-Sq  R-Sq(adj)    AIC
  85.47%     82.28%  28.23


Coefficients

Term          Coef   SE Coef    VIF
Constant    -16.90      4.55
Kills        0.742     0.481   1.25
Damage    -0.01762   0.00435  28.49
Assisted  -0.00862   0.00219   4.07
Credits   0.001196  0.000291  29.47


Odds Ratios for Continuous Predictors

          Odds Ratio       95% CI
Kills         2.1006  (0.8184, 5.3919)
Damage        0.9825  (0.9742, 0.9909)
Assisted      0.9914  (0.9872, 0.9957)
Credits       1.0012  (1.0006, 1.0018)


Regression Equation

P(1)  =  exp(Y')/(1 + exp(Y'))


Y' = -16.90 + 0.742 Kills - 0.01762 Damage - 0.00862 Assisted + 0.001196 Credits
```

The VIF scores are high for Damage and Credits, at 28.49 and 29.47, respectively. Though R-squared, and therefore VIF scores are approximate, this is greater than both values pointed out by the rule of thumb for severe collinearity, which are 10 and 1/(1-.85) = 6.882. Because of this finding, I will consider other models in a moment.

Coefficients: W/L ~ Kills + Damage + Assisted + Credits

The constant has a coefficient of -16.90. This corresponds to an expected odds of victory of e^(-16.9) which in turn corresponds to a predicted probability of victory very close to zero when I have gotten no kills, done no damage or assisted damage, and received no credits. No such instance is observed in the dataset because all tanks are awarded a base amount of credits just for joining a battle, so interpreting the constant is meaningless.

The odds ratio for kills is 2.1006, which means an additional kill, holding damage, assisted damage, and credits awarded constant, multiplies the odds of victory of 2.1006. The coefficient for kills is possibly not statistically significant, as shown above, but this odds ratio interpretation suggests that executing an enemy tank and removing the possibility that it does any more damage to my tank or my allies' tanks is associated with greatly improved odds of victory. A common tactic in World of Tanks and many other team games is to focus the entire team's efforts on eliminating one threat at a time, while conversely team members try to distribute damage taken more evenly so the team can maintain the same level of damage output for as long as possible, so being able to finish off an enemy can also imply more coordinated allies or disorganized enemies.

The odds ratio for damage is 0.9825, which means an additional point of damage, holding kills, assisted damage, and credits awarded constant, decreases the odds of victory by (1-0.9825 = 0.0175 = 1.75 percent. This odds ratio is difficult to interpret because credits awarded was shown largely to be a function of damage dealt. To increase the amount of damage done while holding the amount of credits awarded constant means that other factors which increase credits awarded had to decrease to compensate. In this light, it seems plausible that increasing the amount of damage dealt would decrease the odds of victory if damage was held constant.

The odds ratio for assisted damage is 0.9914, which means an additional point of assisted damage, holding kills, damage, and credits awarded constant, decreases the odds of victory by (1-0.9914 = 0.0086 = 0.86 percent. Like the odds ratio for damage, interpreting the odds ratio for assisted damage is difficult. However, I notice that this decrease is close to half the corresponding decrease for damage. This is interesting to me because the formula for credits awards assisted damage exactly equal to half that of normal damage. When taken in context with the high VIF scores, it is likely that the credits variable is throwing all of the coefficients off.

Goodness-of-Fit: W/L ~ Kills + Damage + Assisted + Credits

Goodness-of-Fit Tests

Test	DF	Chi-Square	P-Value
Deviance	87	18.23	1.000
Pearson	87	220.40	0.000
Hosmer-Lemeshow	8	20.69	0.008

Observed and Expected Frequencies for Hosmer-Lemeshow Test

	Event Probability	W/L = 1		W/L = 0	
Group	Range	Observed	Expected	Observed	Expected
1	(0.000, 0.003)	0	0.0	9	9.0
2	(0.003, 0.007)	0	0.0	9	9.0
3	(0.007, 0.023)	0	0.1	9	8.9
4	(0.023, 0.088)	0	0.4	9	8.6
5	(0.088, 0.956)	8	6.8	2	3.2
6	(0.956, 0.984)	9	8.8	0	0.2
7	(0.984, 0.992)	9	8.9	0	0.1
8	(0.992, 0.995)	8	9.0	1	0.0
9	(0.995, 0.999)	9	9.0	0	0.0
10	(0.999, 1.000)	10	10.0	0	0.0

The Hosmer-Lemeshow test which is the only applicable goodness-of-fit test of the three in this binary response case has a p-value of 0.008, which strongly suggests that the data in their current arrangement is not allowing the logit function to fit properly. Again, this could be due to collinearity problems.

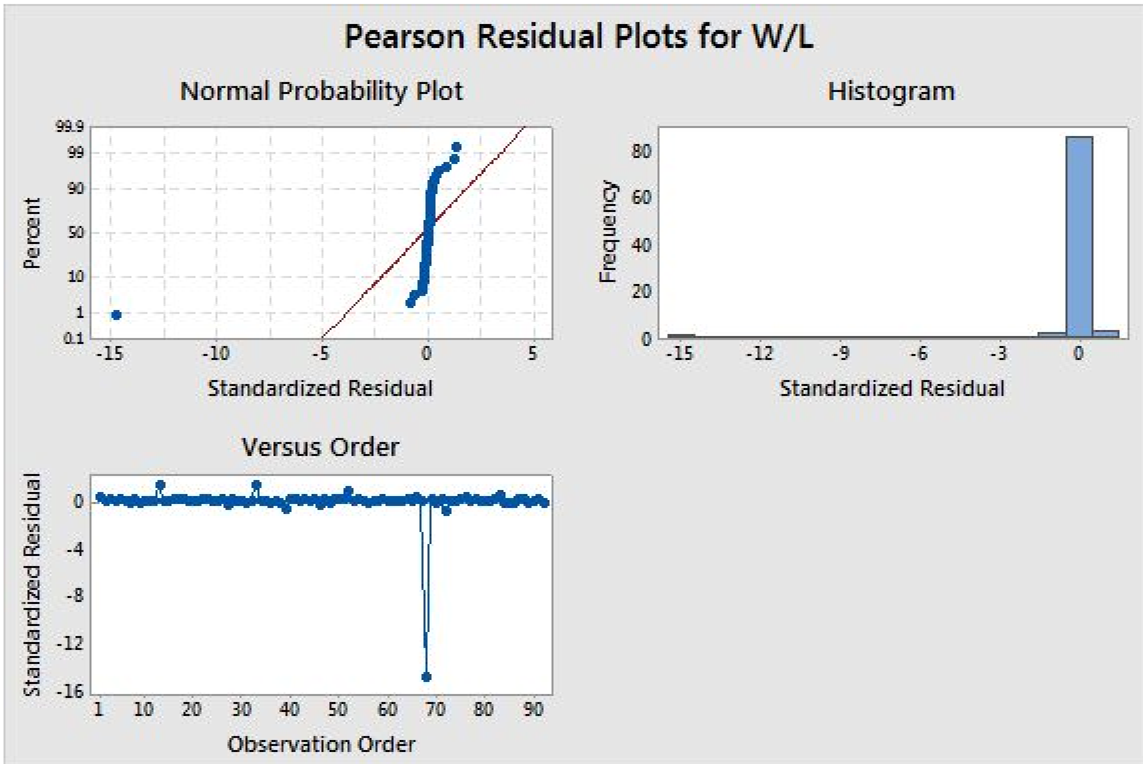
<u>Measures of Association: W/L ~ Kills + Damage + Assisted + Credits</u>

Measures of Association

Pairs	Number	Percent	Summary Measures	Value
Concordant	2032	98.3	Somers' D	0.97
Discordant	27	1.3	Goodman-Kruskal Gamma	0.97
Ties	8	0.4	Kendall's Tau-a	0.48
Total	2067	100.0		

Association is between the response variable and predicted probabilities

Somer's D has a value of 0.97 which is surprisingly high to me. Each team is composed of fifteen players, and a single player will find it difficult to decide the battle no matter how well he or she does. There are many games where I do poorly and win, and also some games where I play well and lose, yet 98.3% of the pairs listed are concordant which means the expected probability of my wins were greater than the expected probability of my losses for 98.3% of the win-loss combinations; a little bit high, since I am just an average player who frequently has no strong effect on the game. I went back to the credits breakdown and found that I am given a very small boost in credits every time I win the game; it's possible that this slight boost is what is causing the predictive power to be so high.



There is one clear outlier, data point 68 with a Pearson standardized residual of -14.7681. This is a battle where I lost after getting 1 kill and dealing 2154 damage and 0 assisted damage. For this battle, I was awarded 49,733 credits. An extremely similar data point is data point 70, where I also lost the battle, also got 1 kill, dealt 2137 damage and 94 assisted damage, and was rewarded 42,725 credits (actually, both battles even took place on the same game map). The credits discrepancy confirmed my suspicions that the model was predicting based on the credits bonus for win/loss. Though I lost the battle in data point 68, I was awarded a “Steel Wall” battle hero model for absorbing large amounts of damage without dying. The forum post says that getting a battle medal means that I get the credits bonus even if I lose, which would explain the credit differential here because I did not get a battle hero medal in data point 70.

Because damage and assisted damage make up the majority of the weighting for the credits formula, adding them into the model along with credits allows the model to discern the constant shift in credits awarded between wins and losses. This allows predicting win or loss to become an almost trivial exercise.

Best Subsets Regression: W/L versus Kills, Damage, Assisted, Credits

Response is W/L

Vars	R-Sq	R-Sq (adj)	R-Sq (pred)	Mallows Cp	S	Kills	Damage	Assisted	Credits
1	26.4	25.6	24.0	302.1	0.42866				X
1	13.9	12.9	10.3	368.2	0.46360	X			
2	47.4	46.2	43.3	193.0	0.36456		X		X
2	28.3	26.7	24.4	293.9	0.42538	X			X
3	82.2	81.6	80.4	10.4	0.21327		X	X	X
3	50.8	49.2	45.7	176.6	0.35431	X	X		X
4	83.6	82.8	81.0	5.0	0.20590	X	X	X	X

Credits is an important variable according to best subsets, since it features in all possible models. Damage is the second most important variable, but damage is also highly correlated with credits. Even though best subsets for logistic regression is only a rule of thumb, the best subsets strongly points to the model with all four variables included as the best one with Mallow's minimized and equal to p+1=5, with highest adjusted and predicted R-squareds.

Dropping damage from the model would probably be the best course of action for eliminating the collinearity while maintaining the most information, but from a practical point of view it would be better to use information that can be obtained before a game is finished to forecast the odds of victory. When I start a game, I'd like to walk in knowing how much damage I need to do to have my game associate itself with a desired probability of victory; knowing how many credits I need to get after the game to associate the game with a desired probability of victory is not very useful. Thus, I will be dropping credits from the model, and rerunning best subsets as a fast check to see if any variable should be dropped.

Best Subsets Regression: W/L versus Kills, Damage, Assisted

Response is W/L

Vars	R-Sq	R-Sq (adj)	R-Sq (pred)	Mallows Cp	S	Kills	Damage	Assisted
1	13.9	12.9	10.3	2.5	0.46360	X		
1	9.6	8.6	6.1	7.0	0.47496		X	
2	16.1	14.2	10.8	2.2	0.46032	X	X	
2	14.4	12.5	8.4	4.0	0.46488	X		X
3	16.3	13.4	8.8	4.0	0.46230	X	X	X

Something interesting to note for this best subsets is the R-squared has dropped like a rock. Mallow's CP is minimized where adjusted R-squared is maximized, at the model with only kills and damage and with assisted dropped.. Mallow's CP is equal to p+1 with the three-variable model. I will evaluate both models.

W/L ~ Kills + Damage + Assisted

Binary Logistic Regression: W/L versus Kills, Damage, Assisted

```
* WARNING * When the data are in the Response/Frequency format, the Residuals versus fits
            plot is unavailable.
```

Method

Link function	Logit
Residuals for diagnostics	Pearson
Rows used	92

Response Information

Variable	Value	Count	
W/L	1	53	(Event)
	0	39	
	Total	92	

Deviance Table

Source	DF	Adj Dev	Adj Mean	Chi-Square	P-Value
Regression	3	16.531	5.5103	16.53	0.001
Kills	1	7.025	7.0251	7.03	0.008
Damage	1	1.806	1.8055	1.81	0.179
Assisted	1	0.131	0.1307	0.13	0.718
Error	88	108.870	1.2372		
Total	91	125.400			

The Assisted variable has an approximate p-value of 0.718 which is not statistically significant at any reasonable level. The Damage variable has an approximate p-value of 0.179, which is also not very significant. On the other hand, the previously marginally significant Kills variable is now extremely significant with an approximate p-value of 0.008.

Model Summary

Deviance R-Sq	Deviance R-Sq(adj)	AIC
13.18%	10.79%	116.87

Coefficients

Term	Coef	SE Coef	VIF
Constant	-1.033	0.494	
Kills	0.599	0.242	1.17
Damage	0.000447	0.000337	1.20
Assisted	0.000166	0.000461	1.04

Odds Ratios for Continuous Predictors

	Odds Ratio	95% CI
Kills	1.8201	(1.1319, 2.9268)
Damage	1.0004	(0.9998, 1.0011)
Assisted	1.0002	(0.9993, 1.0011)

The AIC score is 116.87 (corrected AIC 117.567), which is much higher than the AIC score of 23.87 of the original model. The VIFs, however, are now close to zero, signifying that the collinearity problem has been nearly eliminated. The R-squared, as in the best subsets, has dropped a lot.

The odds ratios for Kills has dropped from 2.1006 to 1.8201, a non-negligible drop but still well within the positive boundary (and so are its 95% confidence intervals). Damage and Assisted now have positive coefficients, which correspond to odds ratios over one. Both these coefficients have high relative standard errors, however, which means that there is a sizable possibility the true coefficients are actually negative (especially for Assisted, whose standard error is almost three times that of the fitted coefficient). The odds ratios for damage and assisted are extremely small, but that is because the units for damage and assisted damage are also small; the Type 59 deals 200-300 damage per shot and has a hitpoint pool of 1300. Dealing 1024 damage, for example, while holding kills and assisted damage constant, is associated with 1.0004^1024 = 1.5061 higher odds of victory. If the prior odds for 1:1, for example, this would be associated with a change to the probability of victory from ½ = 50% to 1.5/2.5 = 60% probability. It is only when damage done reaches 2000 or more that there is a sizable association with higher odds of victory. This makes sense since the Type 59 has 1300 hitpoints and should, as a rule of thumb, be expected to deal around that much on average to maintain a 50% winrate.

```
Regression Equation

P(1)  =  exp(Y')/(1 + exp(Y'))


Y' = -1.033 + 0.599 Kills + 0.000447 Damage + 0.000166 Assisted


Goodness-of-Fit Tests

Test             DF  Chi-Square  P-Value
Deviance         88      108.87    0.065
Pearson          88       92.67    0.346
Hosmer-Lemeshow   8        5.13    0.744
```

The Hosmer-Lemeshow test now has a p-value of 0.744; all traces of a lack of goodness-of-fit have disappeared with the dropping of the credits variable.

Observed and Expected Frequencies for Hosmer-Lemeshow Test

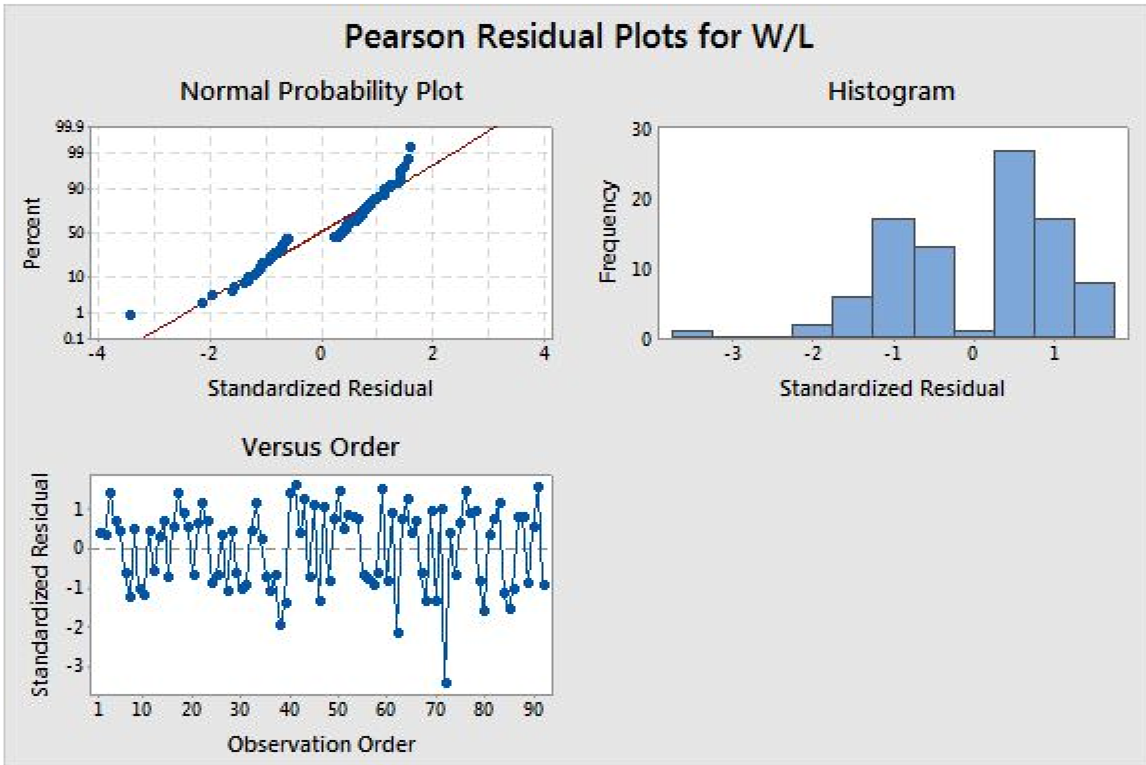
Group	Event Probability Range	W/L = 1 Observed	W/L = 1 Expected	W/L = 0 Observed	W/L = 0 Expected
1	(0.000, 0.315)	2	2.6	7	6.4
2	(0.315, 0.343)	4	3.0	5	6.0
3	(0.343, 0.410)	4	3.5	5	5.5
4	(0.410, 0.476)	4	4.1	5	4.9
5	(0.476, 0.575)	4	5.3	6	4.7
6	(0.575, 0.631)	4	5.5	5	3.5
7	(0.631, 0.697)	8	5.9	1	3.1
8	(0.697, 0.788)	6	6.6	3	2.4
9	(0.788, 0.856)	8	7.5	1	1.5
10	(0.856, 0.955)	9	9.0	1	1.0

Measures of Association

Pairs	Number	Percent	Summary Measures	Value
Concordant	1533	74.2	Somers' D	0.49
Discordant	522	25.3	Goodman-Kruskal Gamma	0.49
Ties	12	0.6	Kendall's Tau-a	0.24
Total	2067	100.0		

Association is between the response variable and predicted probabilities

Somer's D has dropped from 0.97 to 0.49. Though the new model has far less predictive power than the first one, it's much more useful from a practical perspective since now I can get an idea of what sort of performance is associated with a given odds of winning.



The outlier on the far left of the normal probability plot is data point 72, while the outlier on the far right is data point 41. Data point 72 was a loss where I did 2169 damage and 571 assisted damage, while data point 41 was a victory where I did 222 damage and 226 assisted damage. These kinds of games happen reasonably often, both in this dataset and the datasets for the other tanks I play, so I don't think there's anything wrong with them.

Data point 72 is the bottommost data point in the Versus Order plot, which shows fairly constant variance. The points above the 0 line show games where I performed poorly but won, while the points below the 0 line show games where I performed well but lost.

There is good evidence for dropping the Assisted variable, since it is not statistically significant in this model at all and since best subsets had Mallows Cp minimized with the Assisted variable dropped.

W/L ~ Kills + Damage

Binary Logistic Regression: W/L versus Kills, Damage

```
* WARNING * When the data are in the Response/Frequency format, the Residuals versus fits
            plot is unavailable.


Method

Link function               Logit
Residuals for diagnostics   Pearson
Rows used                   92


Response Information

Variable  Value  Count
W/L       1         53  (Event)
          0         39
          Total     92


Deviance Table

Source      DF  Adj Dev  Adj Mean  Chi-Square  P-Value
Regression   2   16.400     8.200       16.40    0.000
  Kills      1    7.100     7.100        7.10    0.008
  Damage     1    2.023     2.023        2.02    0.155
Error       89  109.000     1.225
Total       91  125.400
```

The approximate p-value for Damage is 0.155, which is not statistically significant at a 0.05 level. I will consider the model where Kills is the only predictor after this.

```
Model Summary

Deviance   Deviance
    R-Sq  R-Sq(adj)     AIC
  13.08%     11.48%  115.00


Coefficients

Term          Coef   SE Coef   VIF
Constant    -0.977     0.467
Kills        0.604     0.243  1.17
Damage    0.000466  0.000333  1.17


Odds Ratios for Continuous Predictors

        Odds Ratio       95% CI
Kills       1.8295  (1.1360, 2.9464)
Damage      1.0005  (0.9998, 1.0011)


Regression Equation

P(1)  =  exp(Y')/(1 + exp(Y'))


Y' = -0.977 + 0.604 Kills + 0.000466 Damage


Goodness-of-Fit Tests

Test             DF  Chi-Square  P-Value
Deviance         89      109.00    0.074
Pearson          89       92.95    0.366
Hosmer-Lemeshow   8        2.71    0.951
```

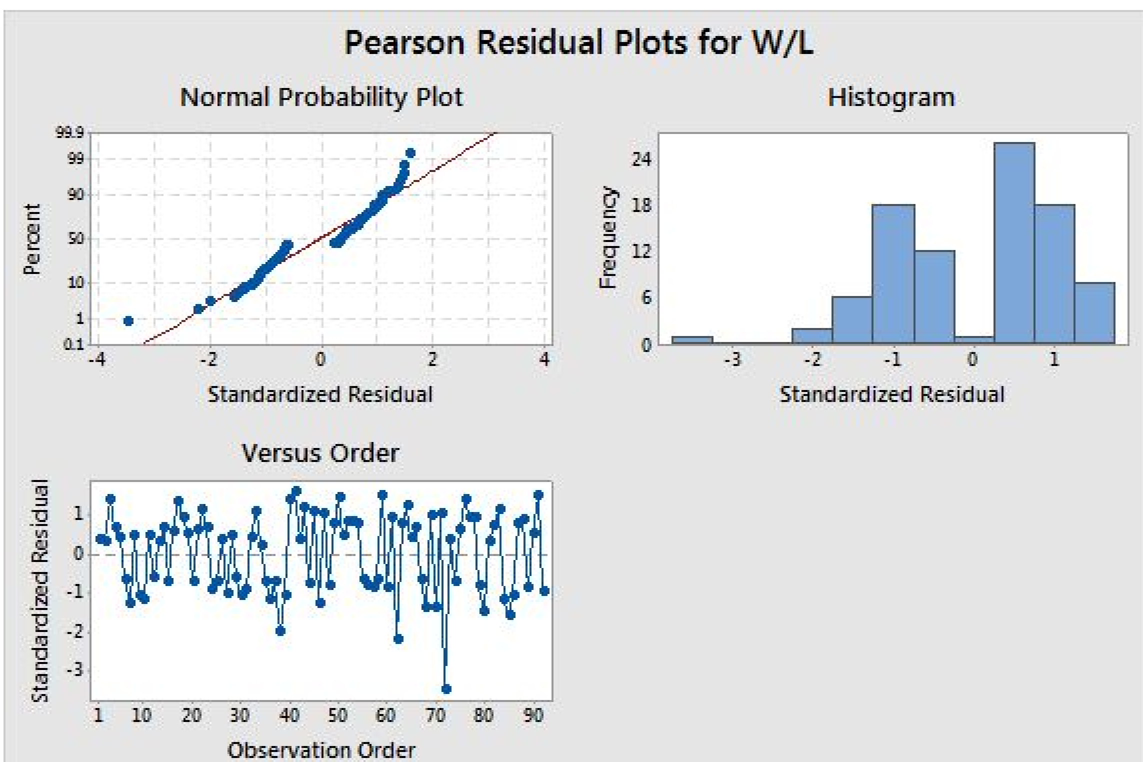
Dropping the Assisted variable does not appear to have degraded the model at all. The AIC score, shown below, has a value of 115 (corrected AIC 115.46), which is slightly lower than the AIC value of the previous model. Additionally, the odds ratios for Kills and Damage are almost the same as the previous model; instead of 1024 damage being associated with 1.0004^(1024)= 1.5061 higher odds of victory, 1024 damage is now associated with 1.0005(^1024) = 1.6684 higher odds of victory. The problem that the coefficient for Damage could still be negative remains. Hosmer-Lemeshow has a chi-squared value of 2.71 which corresponds to an extremely non-significant p-value of 0.951.

Measures of Association

Pairs	Number	Percent	Summary Measures	Value
Concordant	1523	73.7	Somers' D	0.48
Discordant	533	25.8	Goodman-Kruskal Gamma	0.48
Ties	11	0.5	Kendall's Tau-a	0.24
Total	2067	100.0		

Association is between the response variable and predicted probabilities

Somer's D has dropped from 0.49 to 0.48, an almost negligible decrease.

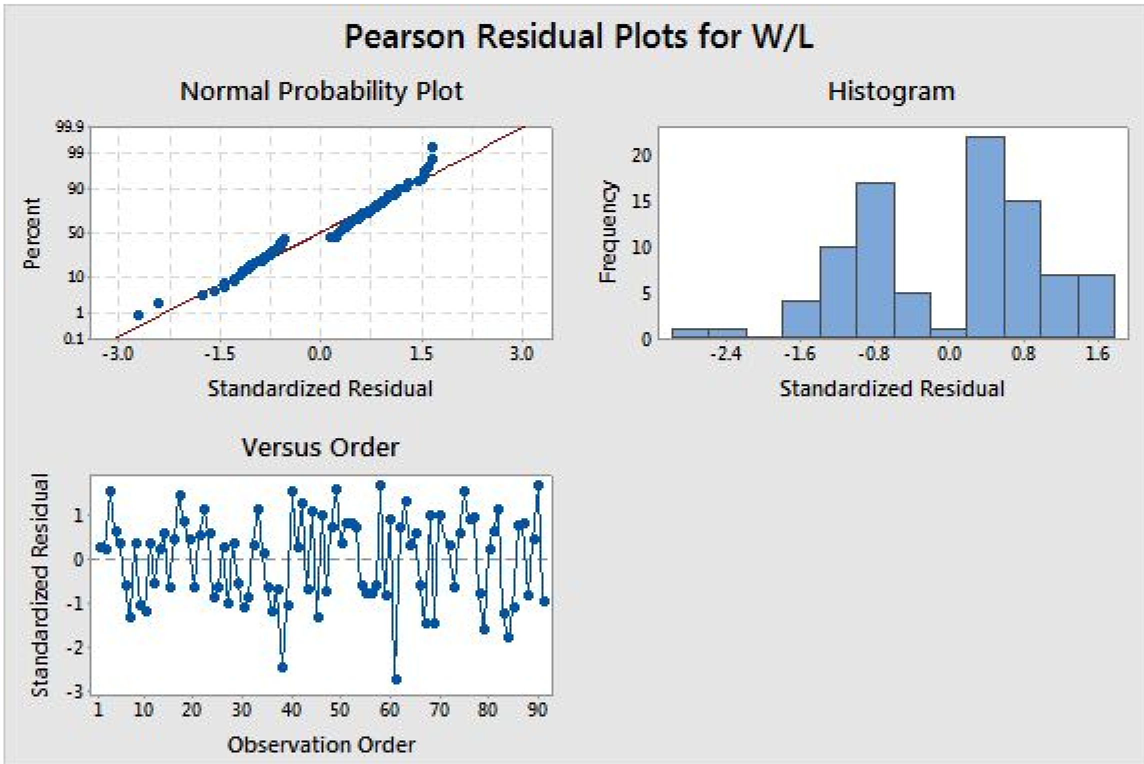


The residuals plots are extremely similar compared to the last model, with data point 72 on the far left of the normal probability plot and data point 41 on the far right. Overall, it appears that the Assisted variable is not a useful predictor of victory and I am not quite sure why this is; high assisted damage variables generally signifies a competent team which can provide good backup and fire support.

The somewhat convoluted explanation I have come up with, which would need further testing to confirm, is that most assisted damage comes from artillery, which focuses on taking out enemy heavy tanks. As a medium tank, the Type 59 generally does not engage with heavies unless it has successfully flanked them, and my personal playstyle is to skirt around the main conflict to hunt down enemy mediums and tank destroyers. This playstyle leads to higher rates of winning than going head to head with enemy heavies, but tends to get less assisted damage. Yet games where I do go up against heavy tanks and win tend to be the ones where I get a lot of assisted damage.

Even though assisted damage is a good thing, the Type 59 is not the kind of tank whose optimal playstyle gets a lot of backup from teammates. Because of this, assisted damage would not be a good predictor of victory just from looking at the data, even though getting assisted damage in the game is always good.

After Removing Outliers



Eliminating data points 72 and 41 did not change much; the odds ratio for Damage is still 1.0005, but the Kills ratio has gone up to 2.2085. The approximate p-value for Damage has dropped to 0.11. The scores for Best Subsets are nearly unchanged as well.

W/L ~ Kills

Deviance Table

Source	DF	Adj Dev	Adj Mean	Chi-Square	P-Value
Regression	1	14.38	14.377	14.38	0.000
Kills	1	14.38	14.377	14.38	0.000
Error	90	111.02	1.234		
Total	91	125.40			

A regression run only on Kills is extremely statistically significant.

Model Summary

Deviance R-Sq	Deviance R-Sq(adj)	AIC
11.46%	10.67%	115.02

The AIC is 115.02, just a little bit larger than the previous model with Kills and Damage. The corrected AIC is 115.293, which is lower than the corrected AIC of the previous model.

Goodness-of-Fit Tests

Test	DF	Chi-Square	P-Value
Deviance	90	111.02	0.066
Pearson	90	95.37	0.329
Hosmer-Lemeshow	3	3.83	0.280

Hosmer-Lemeshow, which was previously 0.95, is now 0.28. This is probably because the Kills variable only takes on 5 values within the entire dataset, making it difficult to fit the logit curve.

Measures of Association

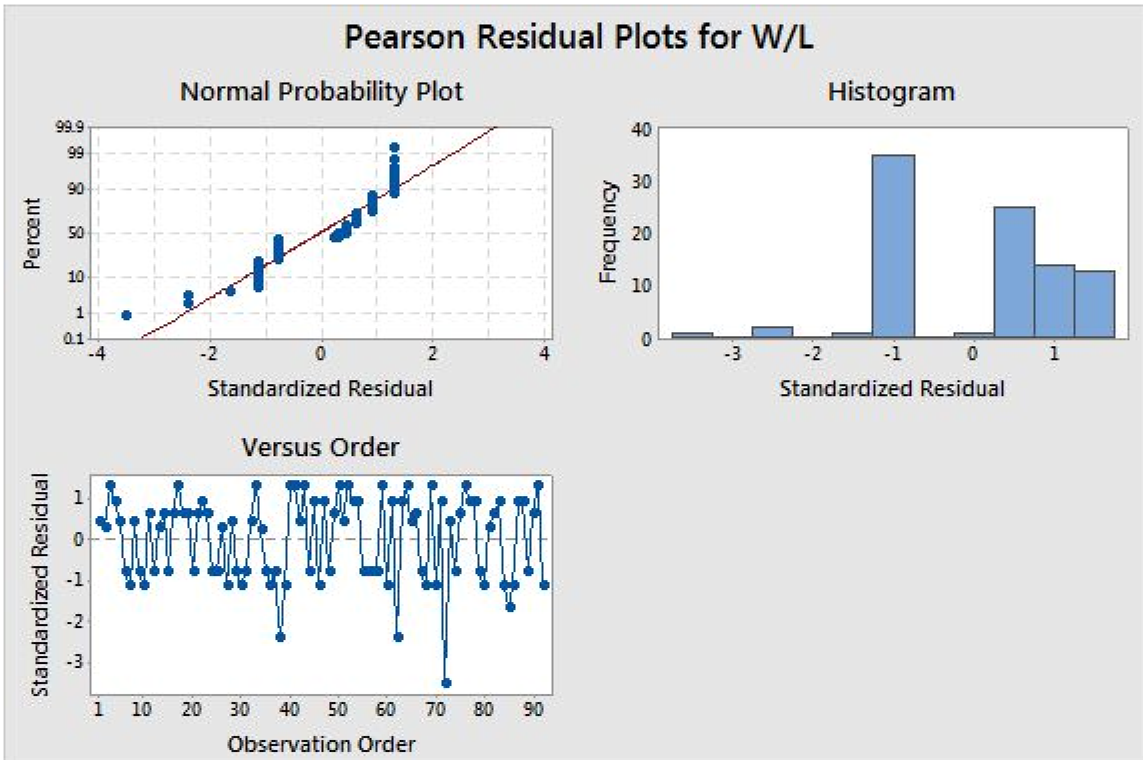
Pairs	Number	Percent	Summary Measures	Value
Concordant	1229	59.5	Somers' D	0.43
Discordant	335	16.2	Goodman-Kruskal Gamma	0.57
Ties	503	24.3	Kendall's Tau-a	0.21
Total	2067	100.0		

Somer's D is now 0.43, which is a notable drop from 0.48, the value for the model with both Kills and Damage.

```
Odds Ratios for Continuous Predictors

        Odds Ratio       95% CI
Kills       2.1036   (1.3552, 3.2652)
```

The odds ratio for kills is now 2.1036, a modest increase over the two-variable model but not really very different. The 95% confidence interval is quite large, from 1.3552 to 3.2652.



The leftmost residual is still data point 72, while the right and topmost is data point 91, another battle where I underperformed and won.

Removing the outliers 72 and 91 did not make the residuals look any better, but did bump the odds ratio for kills up to 2.5741. The AIC went down to 107.29, Somer's up to .48, and Hosmer-Lemeshow has a p-value of 0.35. Because I do not believe these outliers are actually that uncommon and because they do not fundamentally change the model coefficients, I will stick with the models where outliers are not removed.

Between the model with Kills and the model with both Kills and Damage, I prefer the second because the model has a lower AIC score, less evidence of a lack of goodness-of-fit, and a modestly better Somer's D score. Both models will be verified using another 27 data points I gathered while doing this homework.

Model Verification

VW/L	Vkills	Vdamage	PFITS-Kills	PFITS - Kills/Damage	
0	4	2281	0.922627	0.924381	LW
0	0	529	0.378484	0.325222	LL
1	2	1391	0.729344	0.706873	WW
1	2	1734	0.729344	0.738894	WW
1	3	1762	0.850043	0.839877	WW
0	1	804	0.561599	0.500605	LW
1	2	499	0.729344	0.614008	WW
0	2	2237	0.729344	0.781570	LW
0	2	1635	0.729344	0.729887	LW
0	0	1273	0.378484	0.405439	LL
1	0	484	0.378484	0.320633	WL
1	0	1245	0.378484	0.402295	WL
1	2	1839	0.729344	0.748231	WW
1	1	338	0.561599	0.446473	WL

0	2	1469	0.729344	0.714354	LW
1	1	731	0.561599	0.492094	WL
1	0	509	0.378484	0.323178	WL
1	3	1649	0.850043	0.832662	WW
1	0	2259	0.378484	0.519251	WW
0	1	1373	0.561599	0.566557	LW
0	0	896	0.378484	0.363849	LL
1	2	1717	0.729344	0.737361	WW
1	1	3425	0.561599	0.772929	WW
1	0	0	0.378484	0.273564	WL
1	0	1523	0.378484	0.433829	WL
1	2	1514	0.729344	0.718617	WW
0	2	1621	0.729344	0.728598	LW

	Predicted for both one and two-variable model	
Actual	Victory	Loss
Victory	10	7
Loss	7	3

The model was only able to correctly classify 13 of the 27 data points, which is incredibly low. This is not any better than flipping a coin, or even taking my base win rate for the Type 59 (51% across 1300 battles) and using that. However, looking at the prediction dataset, there were a lot of anomalous battles, such as the first data point where I got 4 kills and did 2281 damage but lost or the fourth-from-last data point where I got no kills and did no damage and still won; these sorts of battles were proportionately less common in the original dataset and I want to gather more data to both improve the model and get a better idea of its prediction capabilities before tossing it out altogether. Moreover, being one teammate among 15, it should be expected that small data samples such as the verification dataset experience high variability.

I subtracted 0.5 from all of the values and took the summation to find the net deviations. The one-variable model had a net deviation of 2.653 while the two-variable model had a net deviation of 2.46 - almost identical. The absolute deviations were even closer, at 4.818 and 4.888 respectively. Because Kills is weighted extremely heavily compared to other predictors, both models predicted almost identically to each other except for victories where i did a lot of damage; in these battles, the two-variable model tended to predict a much higher probability of winning, while in the battles where I did a lot of damage, but lost, the two-variable model was off, but only by a tiny bit. Using the three-variable model with assisted damage included to predict victory also yielded nearly identical results to the other two.

Overall, all the models agree with each other, though their abilities to predict victory are still in question. Contrary to my expectations, Kills on its own seems to be a very good standalone predictor of victory rate; the model only predicts I am more likely to lose than to win when I get 0 kills. The effects of any damage or assisted damage I do do not heavily affect my odds of winning unless I do very well. For other tank types, such as light tanks or artillery, the effect of assisted damage and damage may be more important, but that would have to be tested another time.