Introduction

A very important part of our history is all the wars that have happened. They have shaped our society to what we know today. A few wars in particular that have had a huge impact include but are not limited to World War II (1941-1945), the Vietnam War (1964-1973), the Korean War (1950-1953), and the Gulf War (1990-1991). Two wars in particular that were included in the first part of our project, but are not included in our final report are World War I (1917-1918) and the War on Terrorism/Afghanistan (2001-Present). The reason for this is because the data was very hard to obtain, if at all obtainable and it was very limited to what we could do. We wanted to broaden our horizons.

Through these wars several questions will be answered including a question about recession, deficit, per capita income, and poverty involving Gross Domestic Product (GDP) and Gross National Product (GNP). Through the analysis of the data obtained through these questions the ultimate question will be answered. Is war profitable and beneficial to the American Economy? Using various topics learned through class we will deduce a null and alternative hypothesis and test this throughout the report. The null hypothesis is war does not affect the GDP and thus does not affect the economy. The alternative hypothesis is that war does have an effect on GDP and thus does affect the economy. These hypothesis will be concluded in our final analysis at the end of the report after all the data is obtained. We want to deduce whether or not war is profitable through GDP to the American Economy. The data that will be collected are combined from three years prior to the specified war, the war itself, and three years after the war to ultimately gain answers to the questions posed above.

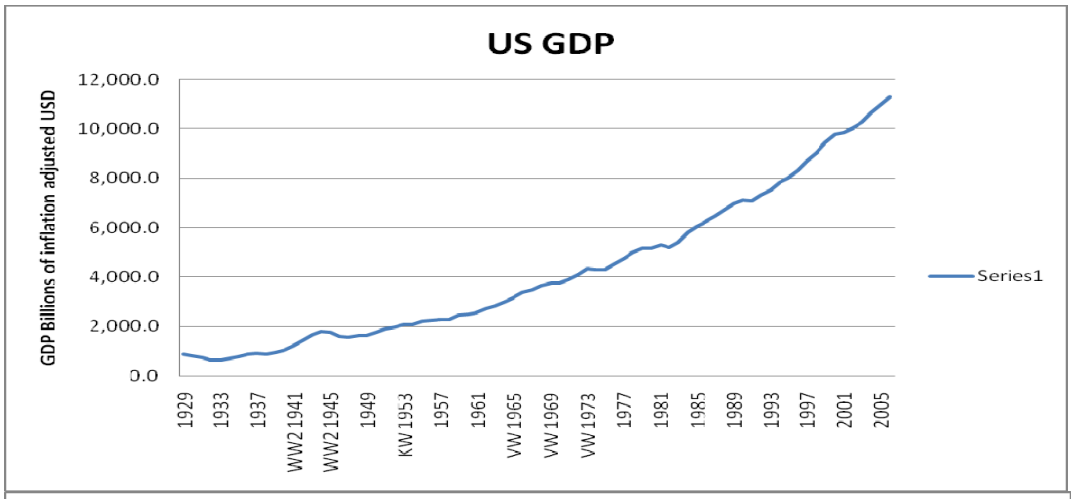
GDP and GNP

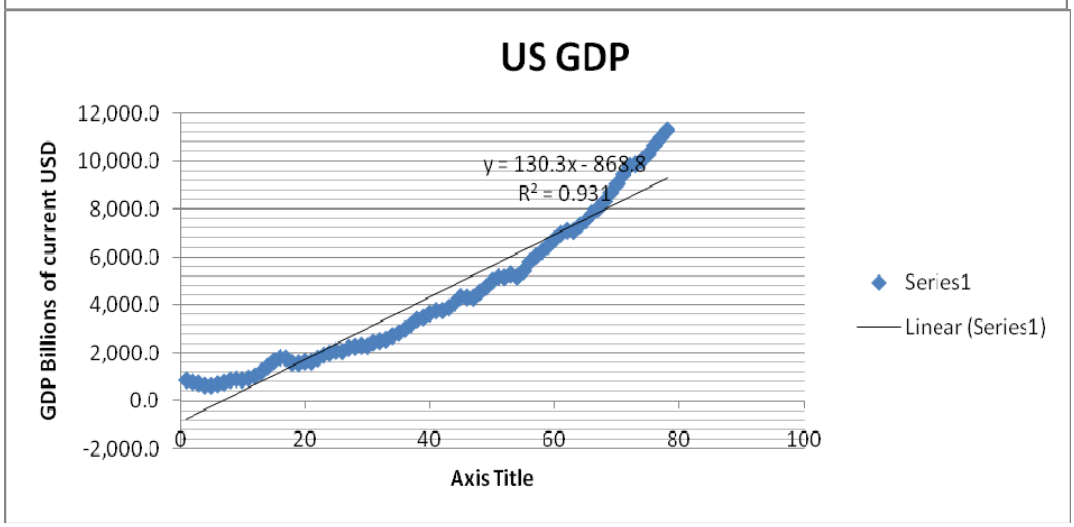
GDP is one of the most comprehensive and closely watched economic statistics: It is used by the White House and Congress to prepare the Federal budget, by the Federal Reserve to formulate monetary policy, by Wall Street as an indicator of economic activity, and by the business community to prepare forecasts of economic performance that provide the basis for production, investment, and employment planning.

GDP is composed of goods and services that are produced for sale in the “market”—the generic term referring to the forum for economic transactions—and of nonmarket goods and services—those that are not sold in the market, such as the defense services provided by the Federal Government, the education services provided by local governments, the emergency housing or health care services provided by nonprofit institutions serving households (such as the Red Cross), and the housing services provided by and for persons who own and live in their home (referred to as “owner-occupants”). However, not all productive activity is included in GDP. Some activities, such as the care of one's own children, unpaid volunteer work for charities, or illegal or black-market activities, are not included because they are difficult to accurately measure and value.

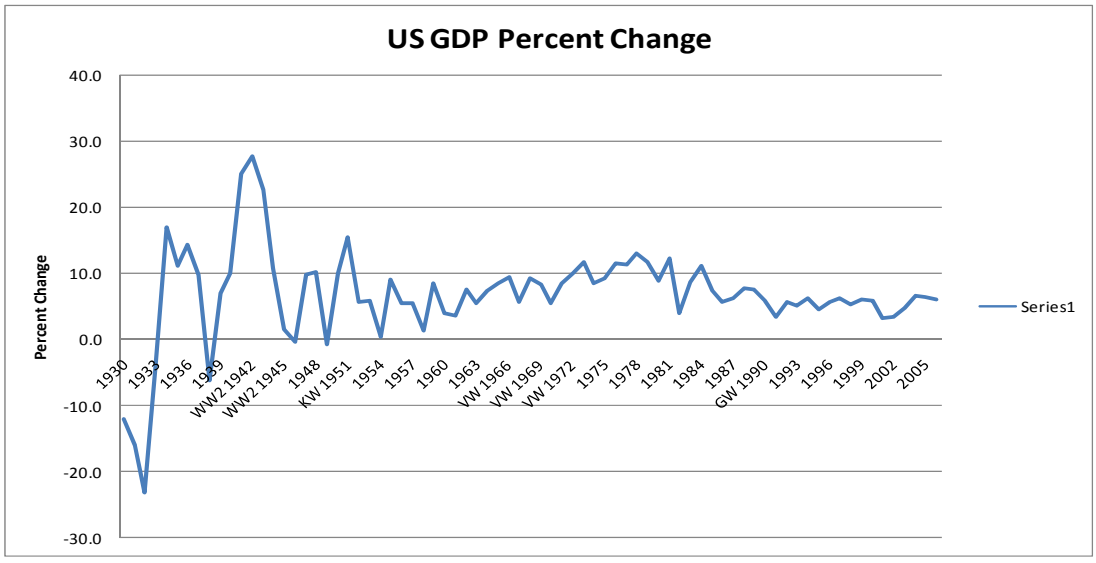
The graphs below have explanations above them but are adjusted for the inflation. This is what we call chained dollars. To figure this out, we used the formula (1+ (Real Growth % - CPI %))*previous years GDP. Consumer Price Index (CPI) is an index number measuring the average price of consumer goods and services purchased by households. The percent change in the CPI is a measure of inflation. This will give us more reasonable data going year by year for a better understanding and realism during the above stated wars. Throughout the report anything using the GDP will be adjusted for inflation and change.

The below graph illustrates how the US GDP has been almost always increasing from the year prior. It seems to grow exponentially.

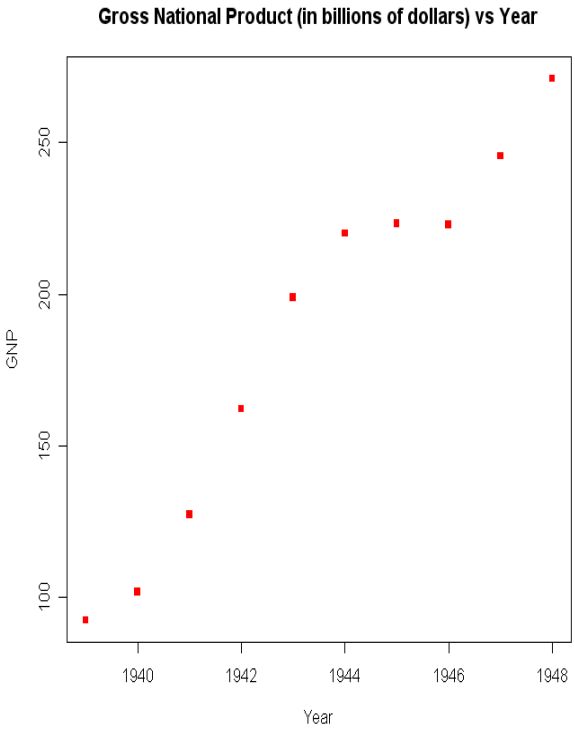


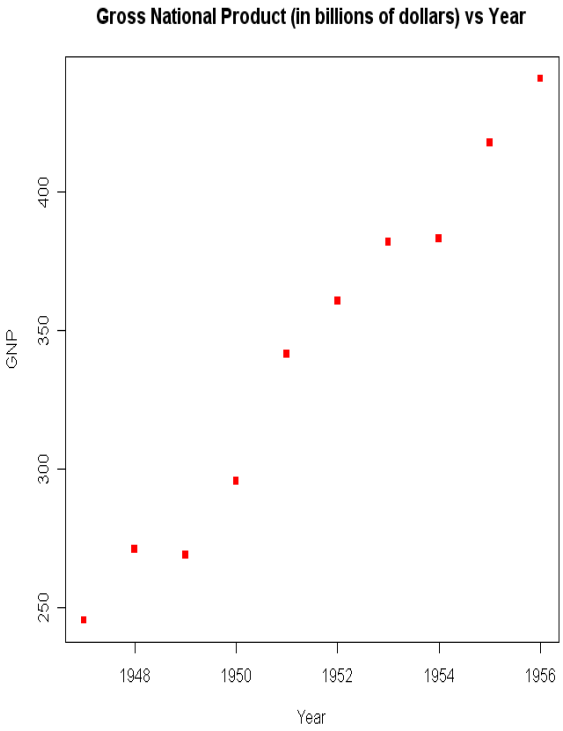


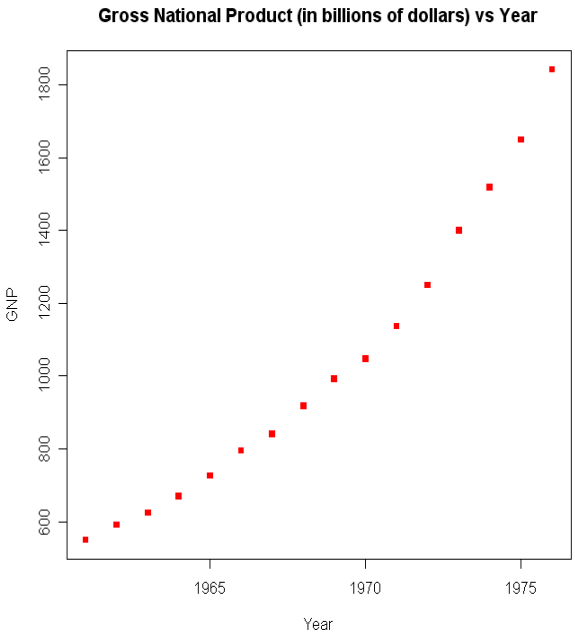
The below graph shows the percent change in US GDP by year. By observing the data one may surmise that wars do indeed impact the GDP by a considerable amount. The graph shows clear spikes in the percent change of the GDP during nearly every major US conflict.

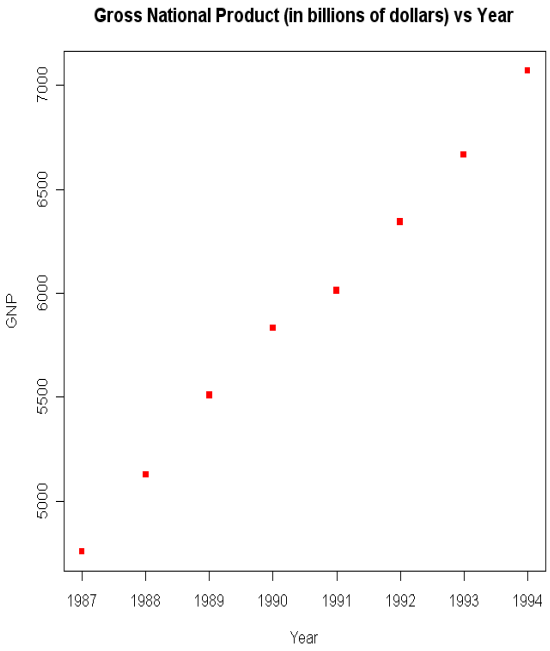


The Gross National Product (GNP) is the total dollar value of all final goods and services produced for consumption in society during a particular time period. Its rise or fall measures economic activity based on the labor and production output within a country. The tangible factors used to calculate this include, but are not limited to cars, furniture, bread, education, healthcare, and auto repair. Below are four graphs in conclusive with the wars showing how GDP increases.







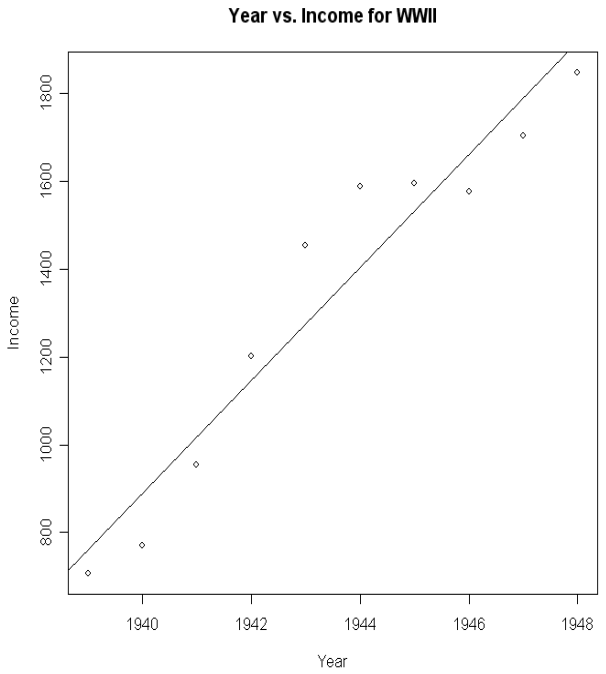


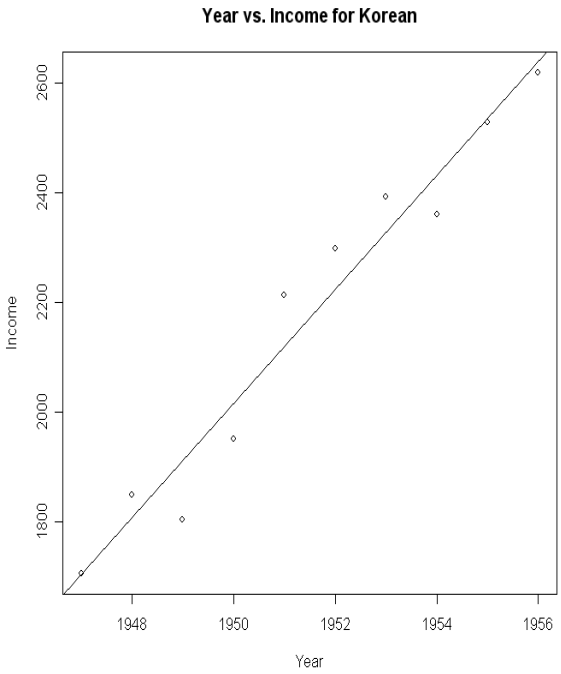
Through these two explanations of GDP and GNP one might ask what the difference between them is. Here is a better explanation coming them so that the difference is justified. A country's GDP is

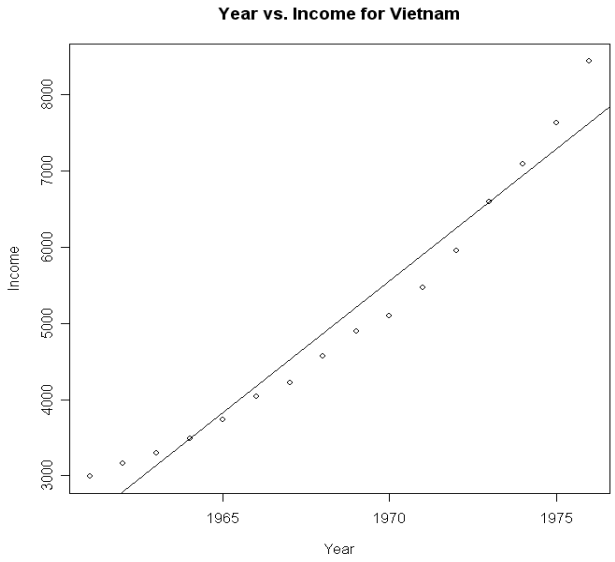
the amount of goods and services, measured as market prices, produced within a country during a particular time period (usually a year). GNP is the amount of goods and services produced by residents of a country, regardless of where that production takes place.

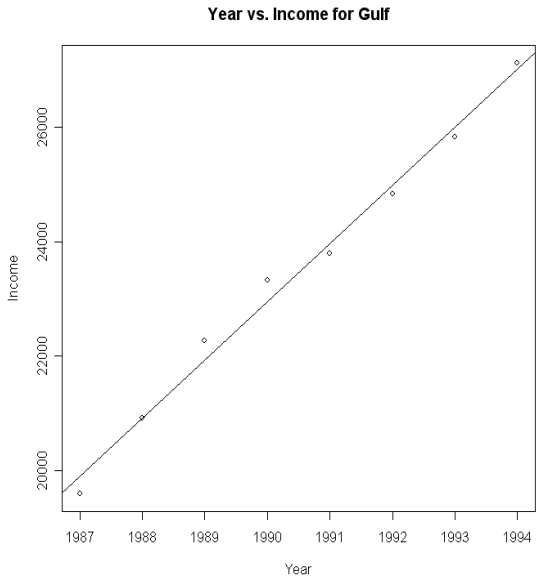
Question: Does War effect per capita income?

We first have to define what per capita income is. It is the mean income computed for every man, woman, and child in a particular group. It is derived by dividing the total income of a particular group by the total population. It reflects gross nation product or GNP. The factors that effect per capita income include but are not limited too wealth as in the euro or the US dollar, national income levels, and personal income for each individual person or home. In all of these wars, you will see that per capita income starts at a low and rises during a war and peaks higher after the war.

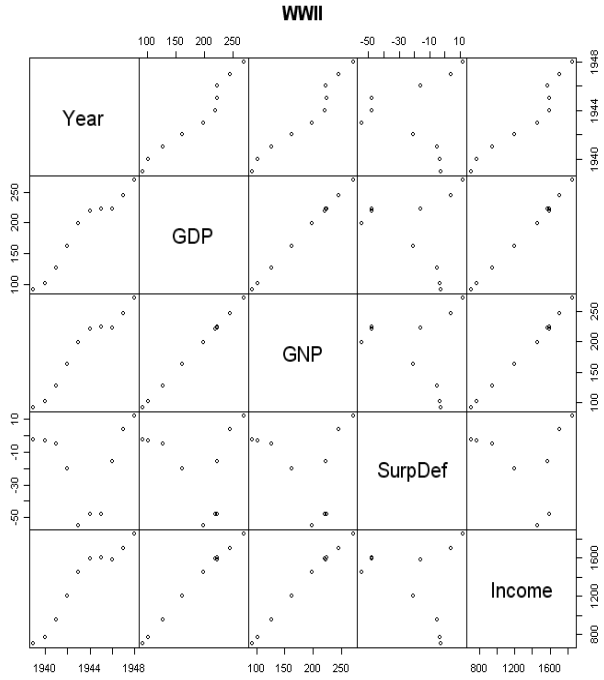


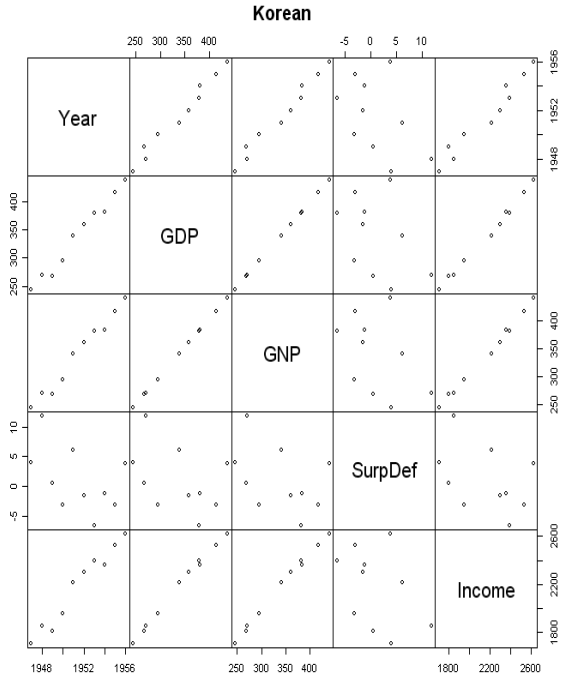


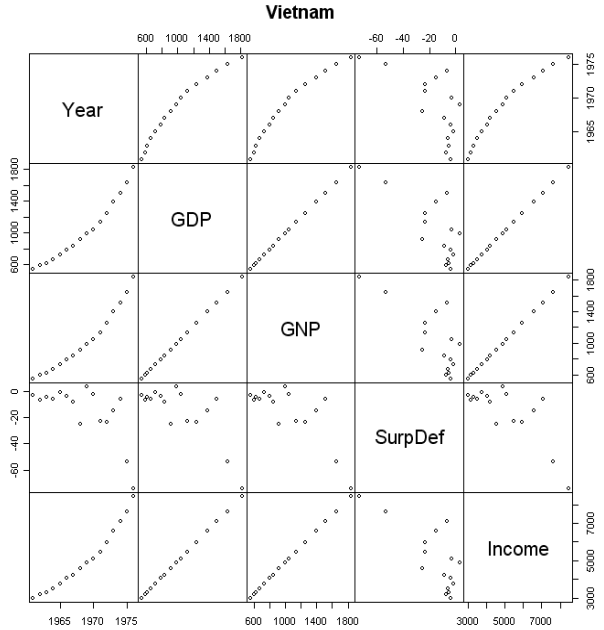


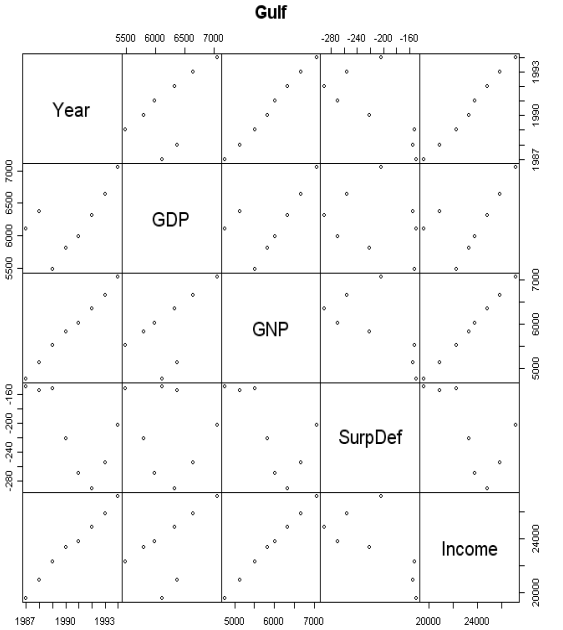


Analyzing these graphs we see that for World War II, the p-value is 1.001*10^-5 which is significant with an F-statistic of 95.67 and an R^2 value of 0.9132.The equation for the income of the war is 128.84x-249053.07. The Vietnam War has a p-value of 1.242*10^-10 which is very significant with an F-statistic of 278.2 and R^2 value of 0.9487. The equation for the income of the war is y=345.92x-675907.97. The Korean War has a p-value of 1.11*10^-6 which is significant with an F-statistic of 171 and an R^2 value of 0.9497. The equation for the income of the war is y=104.2x-201200. Finally, we see that the Gulf War has a p-value of 3.445*10^-7 with an F-statistic of 575.6 and an R^2 value of 0.988. The equation for the income of the war is y=1018x-2003000. Through this analysis we see that all of these wars are very significant and effect per capita income greatly. More analysis is to follow.

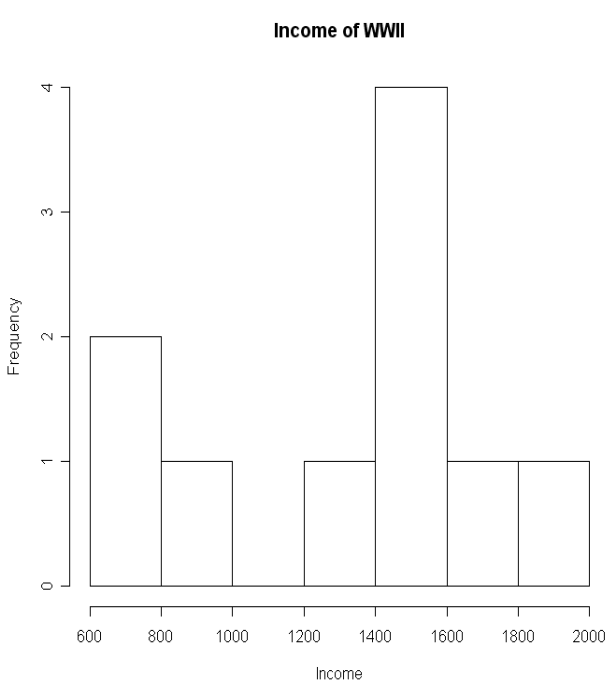


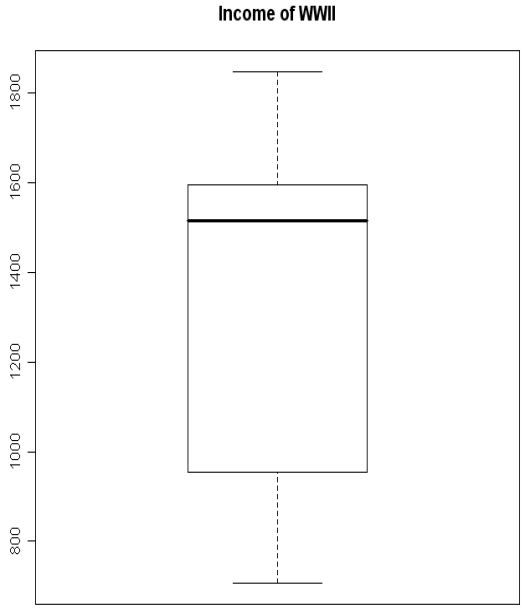






These graphs are deducing multiple regression analysis using GDP, GNP, Surp/Def, and Income. This helps us better understand that per capita income increases with war. Analyzing these graphs we see that for World War II the p-value is 1.217*10^-5 which is significant with an F statistic of 189.1 and R^2 value of 0.9882. The equation for the income for this war with all the factors is y=1937.11744+3.33186(GDP) +0.04650(Surp/Def)-3.14013(Income)-0.01970(Income). In the Korean War, the p-value is 1.262*10^-5 which is significant with an F statistic of 186.4 and an R^2 value of 0.988. The equation for the income for this war with all the factors is y=1944+0.3896(GDP)-0.2757(GNP)-0.02924(Surp/Def)-0.01411(Income). In the Vietnam War, the p value is 4.387*10^-10 which is very significant, with an F-statistic of 191.2 and an R^2 value of 0.9807. The equation for the income for this war is y=1946.67960+0.5558(GDP) +0.4027(Surp/Def) +0.01087(Income)-0.58261(GNP). Finally, for the Gulf War the p-value is 2.469*10^-5 with an F-statistic of 1628 and an R^2 value of 0.9989. The equation of income for this war is y=1979+0.008135(GNP)-5.357*10^-5(GDP)-0.002853(Surp/Def)-0.001585(Income). All of these factors are important in the analysis because income affects the world and it is how we live. We need all of these factors to deduce what income would be. These equations all differ and because of each of these wars we see that even though some part of the equation is negative, the overall is positive for the different values and it shows a positive correlation and income increasing. Below is some more graphical analysis.

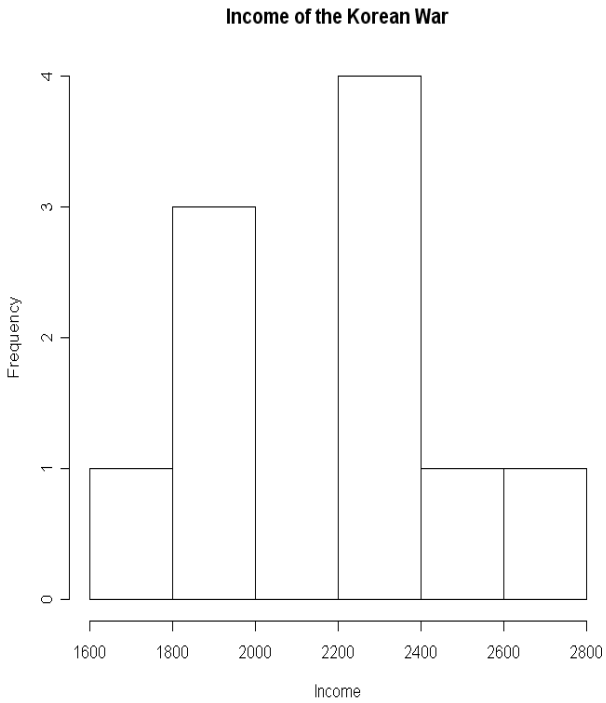


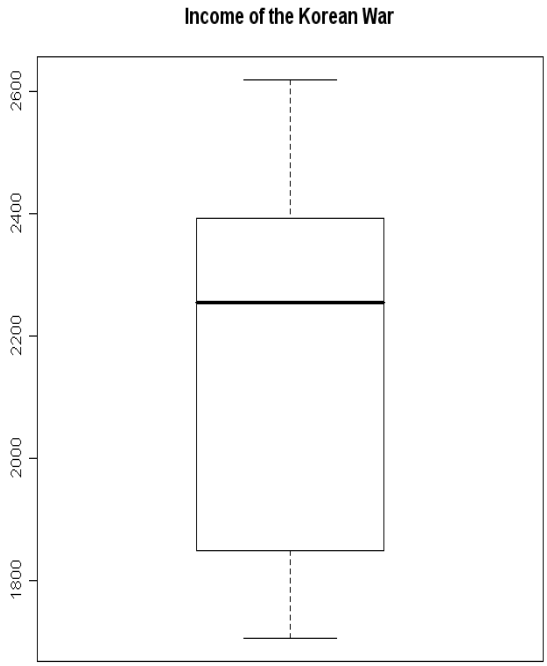


This shows a Histogram and Boxplot of WWII, which shows slightly left scew-ness and the boxplot shows a range of incomes where the residuals are

Residuals:

Min	1Q	Median	3Q	Max
-119.47	-81.91	-59.47	61.47	185.18

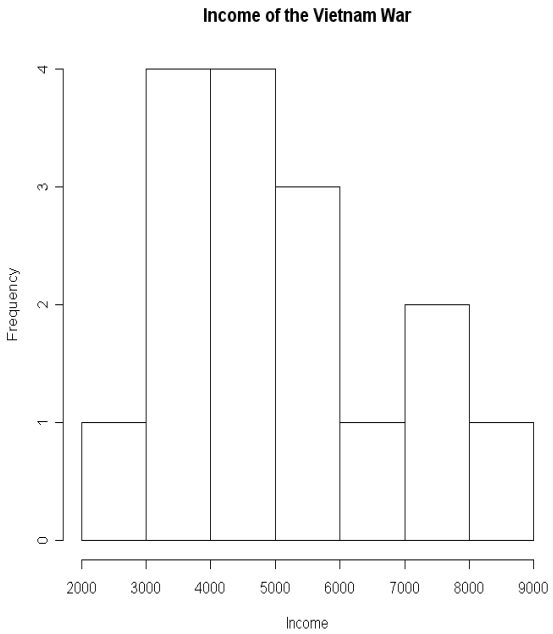


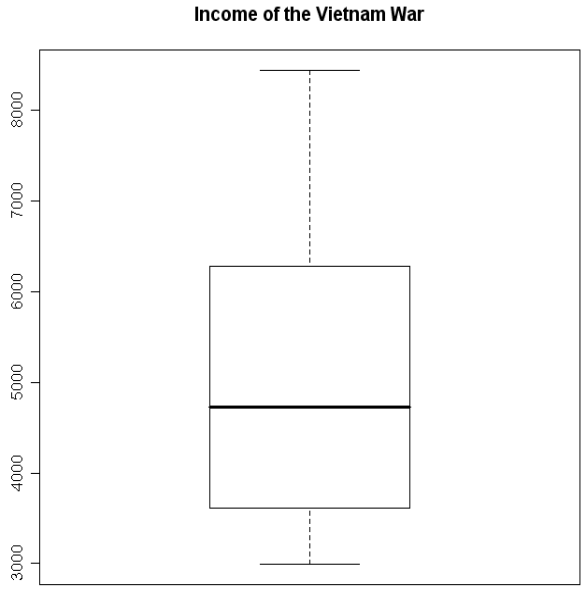


This shows a histogram and boxplot of the Korean War. The histogram is fairly symmetric and the boxplot shows the range of incomes. The residuals for the income in this war are

Residuals:

Min	1Q	Median	3Q	Max
-108.255	-54.300	-3.691	59.145	93.309

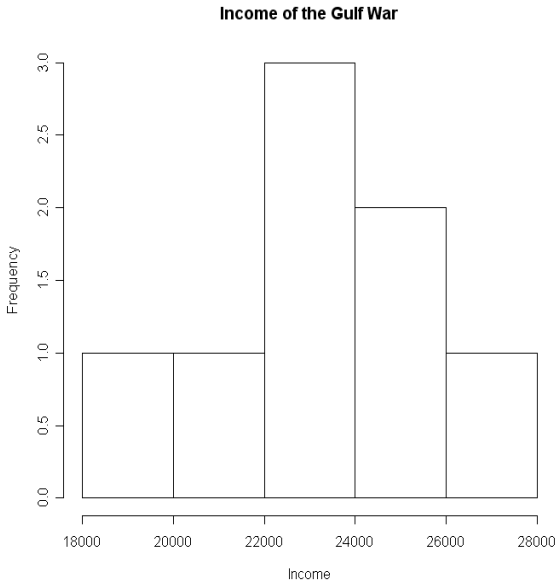


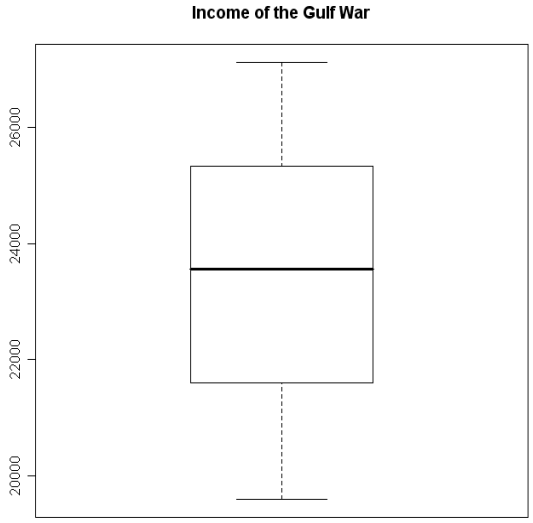


This shows a histogram and boxplot of the Vietnam War. The histogram has a slight skew-ness to the right and the boxplot has a slightly lower increase in incomes, but none the less still the variety of incomes. The residuals are:

Residuals:

Min	1Q	Median	3Q	Max
-460.39	-300.52	-50.46	199.31	805.07

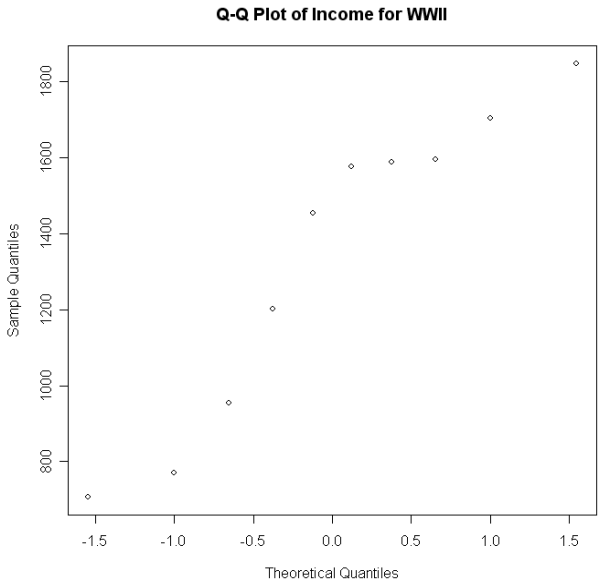


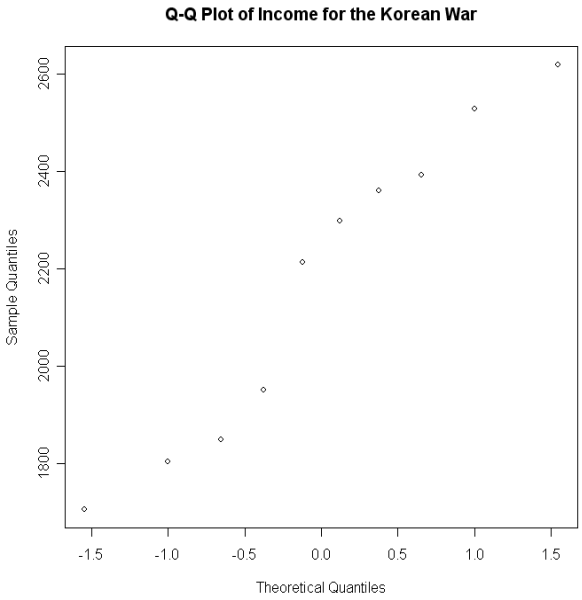


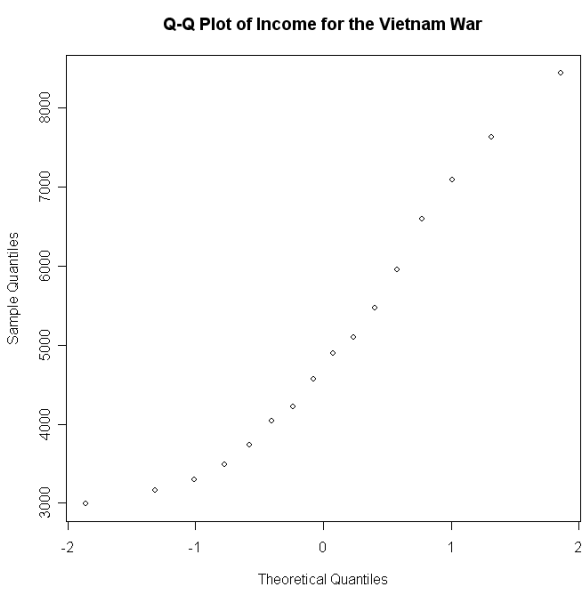
This shows a histogram and boxplot of the Gulf War. The histogram has a slight, very slight, if any at all, skew-ness to the left and the boxplot shows the variety of incomes basically evenly throughout the chart. The residuals are:

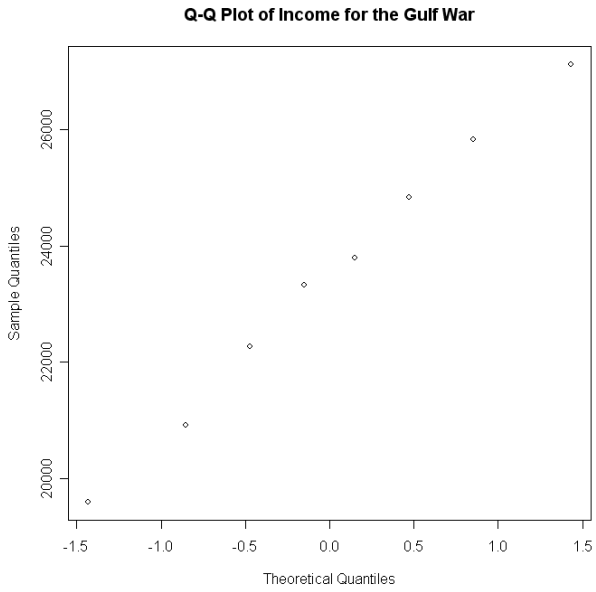
Residuals:

Min	1Q	Median	3Q	Max
-312.5	-178.3	-75.4	164.7	378.6









All four of these Q-Q Plots show increasing income going along the lines of each war. Some are more than others, but overall there is a steadily increasing factor. The factor is that war creates jobs and through these jobs, more citizens get paid and income increases. With income increasing, profit increases and thus confirming our thoughts that per capita income increases and is profitable to the American Economy.

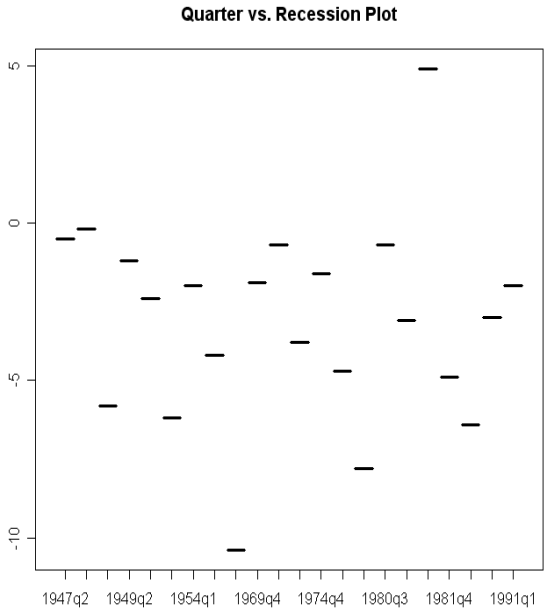
Finally, we will do some mathematical analysis for each war. In World War II, the mean income is 1340.4 with a standard deviation 406.05287. The 95% confidence interval for this war is 1088.730 to 1592.070. In the Korean War, the mean income is 2171.8 with a standard deviation of 322.83116. The 95% confidence interval for this war is 1971.711 to 2371.889. In the Vietnam War, the mean income is 5042.5 with standard deviation 1687.86358. The 95% confidence interval for this war is 4215.462 to 5869.538. Finally, in the Gulf War, the mean income is 23461.375 with standard deviation 2506.4556. The 95% confidence interval is 21724.52 to 25198.23. These confidence intervals are very in sync and can be sure that the level of accuracy is really close.

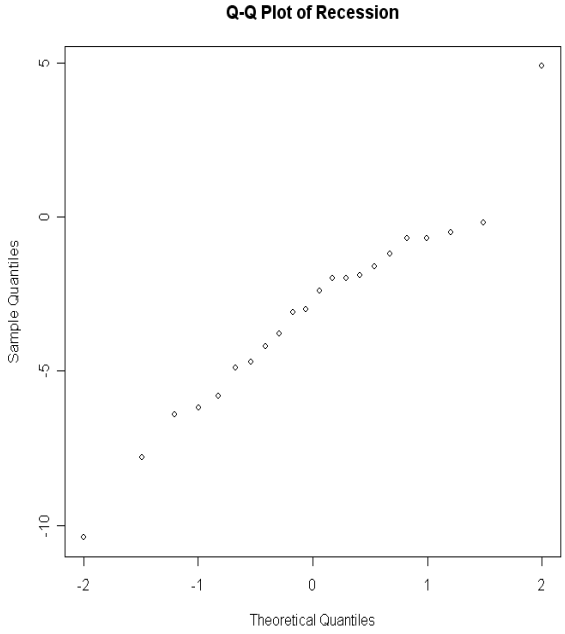
Question: Did War make us go into recession or pull us out of recession?

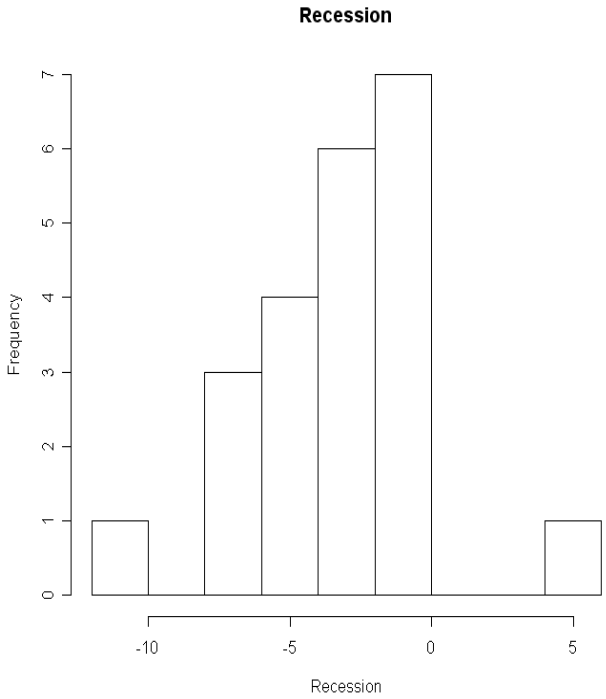
A recession is a decline in any country's gross domestic product (GDP), or negative real economic growth, for two or more successive quarters of a year. It is important for the report to consider at what times the US has entered recessions. If they seem to be common before, during, or after wars there may be a direct link between the two. This will ultimately help with the final conclusion of the report which is to fully support, or not, that wars economically benefit the economy of the United States of America.

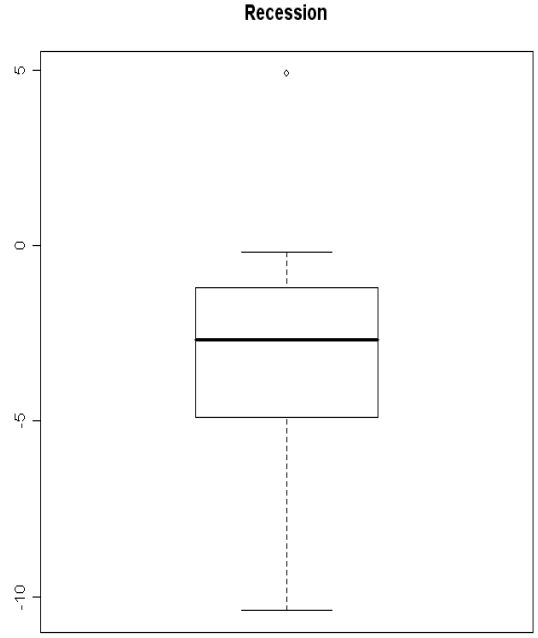
The United States has faced very few recessions. In total there are 22 quarters of the 304 quarters of data that were available. This is only about 7.24% of the quarters the US was in a recession. Below is some graphical analysis with some data analysis and a list of all the dates and periods of recessions.

1947q2	-0.5	1947q3	-0.2	1949q1	-5.8	1949q2	-1.2	1953q3	-2.4
1953q4	-6.2	1954q1	-2.0	1957q4	-4.2	1958q1	-10.4	1969q4	-1.9
1970q1	-0.7	1974q3	-3.8	1974q4	-1.6	1975q1	-4.7	1980q2	-7.8
1980q3	-0.7	1981q2	-3.1	1981q3	4.9	1981q4	-4.9	1982q1	-6.4
1990q4	-3.0	1991q1	-2.0						









The data show that the Korean War seems to have brought us out of a recession (1949) and then dumped us back into another one when it ended (1953 into 54). There was a recession for three quarters after the Vietnam War ended (1974q3 – 1975q1). There was also a recession during the most part of the Gulf War (1990 – 1991). That is a total of 8 quarters of the 22 are related to American wars. This is not counting the recessions that the Korean War may or may not have helped America exit out of. That is a

36.36% of the American recessions are related to American wars. It is quite possible to theorize that war can bring a country out of a recession but also cause it to fall into economic recession once it has ended. An explanation of the many quarters in recession in the early 1980s could be the tight monetary policy in the U.S. to control inflation and sharp correction to overproduction of the previous decade which had been masked by inflation. Doing some mathematical analysis we see that the mean is -3.11818 with standard deviation 3.1741. The 95% confidence interval for recession is -4.444528 to -1.791832 which is a very good approximation and all the residuals are 0 with no residual degrees of freedom.

For this part, we think the amount of quarters the US was in recessions was really good considering the wars, the cost of living, and just life itself. They also were not terrible recessions and we were able to get ourselves out of them rather quickly. In conclusions, these recessions during the war made us prosper and through these we examined great strength, got ourselves out of them and ultimately used this force to make war profitable.

Question: Does war affect Poverty?

When trying to find data of the poverty percentages in this country from before 1959 for Hispanics, Asians, Blacks, and Whites (Caucasians), it was impossible. The Federal Government never tabulated official poverty statistics in the census for public domain use. The poverty analysis that will be done will be in the time period of the Vietnam War, the time in between the Vietnam War, and the Gulf War itself. Data on the poverty levels of different races takes effect late such as for the Gulf War, only Asian and Hispanic poverty levels exist by federal government tabulation at this point. Before the Gulf War, poverty levels for Asians did not exist. However, in the limited data that we do have, we will go ahead with the statistical analysis.

The first analysis that needs to be done is to compare the poverty level of blacks and whites in the Vietnam War, the time in between the Vietnam War to the Gulf War, and the Gulf War itself. The data spans from 1966 to 1994, the end of the time period in question for the whole range of data. In the Vietnam War, a stepwise regression was performed with the national poverty level of all races put together as the response variable. The poverty level for a particular race was considered to be an explanatory variable for the model. Considering a linear regression model with the appropriate values being response and explanatory respectively, we get that the regression model has a multiple R-squared value of .9754, which means that 97.54% of the variation in the poverty level on the national level is explained by the changes in the poverty level of Caucasians and African-Americans during the Vietnam War. The p-value for the significance test is $5.709 \text{ X } 10^{-8}$ which is very small, which is further evidence that the poverty level on the national level of all races is strongly influenced by the poverty levels of Caucasians and African-Americans. The poverty level of Whites as its own explanatory variable in the model has a P-value of $1.12 \text{ X } 10^{-7}$, which shows that the poverty level of Whites during the time of the Vietnam War had a significant effect in determining the overall poverty level of the country. The P-value corresponding to the African-American poverty level was .736, which clearly is not statistically significant. By doing a stepwise regression, we conclude that the best model to predict the poverty level of the country is solely based on knowing the poverty level of Caucasians during the period of the Vietnam War. The regression equation is All Races Poverty Level = .73712+1.21383x (where x is the poverty level of Caucasians.) With the new model, again 97.51% of the variation in overall poverty level is explained by the change in poverty level. The P-value for this new model is $2.379 \text{ X } 10^{-9}$. Hence, in conclusion, during the Vietnam War, the poverty level of the majority race (the Caucasians) was the driving force in explaining the poverty level of the county of the United States in this pivotal war.

Now, onto explain the racial interactions for when we analyze the poverty that happens during peacetime between the Vietnam War and the Gulf War. We now are going to include the poverty levels for Hispanic and see if that has any effect into the model as far as predicting the poverty level overall regardless of race in this country. Again, if one does a regression model to start with and uses all three explanatory variables, the R-squared value is .9996. This model is very good to begin with; this model states that 99.96% of the national poverty level for all races put together is explained by the changes in the poverty levels of the Caucasians, African-Americans, and Hispanics. The equation of the regression

line is All Races Poverty Level = .009684 + 1.005638*White + .1041087*Black + -.022477* Hispanic. What is interesting is this means that the national poverty level In this country is interestingly negatively correlated with the level of Hispanic Poverty level. The P-value overall for the whole model, without doing stepwise regression) is $6.186 \text{ X } 10^{-9}$. Again, this P-value is small showing statistical significance, and thus the national poverty level is strongly influenced by the changes in poverty in Caucasians, African-Americans, and Hispanics. Below is the output for the summary call on the regression equation.

```
> summary(allraces.reg)
Call:
lm(formula = AllRaces ~ White + Black + Hispanic, data = AllRaces.df)
Residuals:
      1       2       3       4       5       6       7       8       9
-0.014339  0.001319  0.033551  0.028351 -0.018230 -0.047495  0.012126 -0.016537  0.021254
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.009684   0.254157   0.038 0.971080
White        1.005638   0.052443  19.176 7.11e-06 ***
Black        0.101487   0.011134   9.115 0.000266 ***
Hispanic    -0.022477   0.018106  -1.241 0.269510
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.03346 on 5 degrees of freedom
Multiple R-Squared: 0.9996,     Adjusted R-squared: 0.9994
F-statistic:  4253 on 3 and 5 DF,  p-value: 6.186e-09
```

As one can see, the P-values associated with the explanatory variables of the Caucasian and African-American poverty level are the most important at play since they are both statistically significant levels.

The goal now is to do a stepwise regression model, and analyze the new regression model in the same way that the full regression model was analyzed. Doing normal stepwise regression, we get the same exact model. The Hispanic Race is included into the regression model. The equation of the regression model still becomes All Races Poverty Level = .009684 + 1.005638*White + .1041087*Black + -.022477* Hispanic

Now, what about the interaction between races does that play into effect into the model, and can it make this model any better. This model is virtually perfect.

By calling the following function: allracesINTERACTION.reg=lm(AllRaces~White*Black+White*Hispanic+Black*Hispanic,data=AllRaces.df)

summary(allracesINTERACTION.reg)→ We get the following output:

```
Call:
lm(formula = AllRaces ~ White * Black + White * Hispanic + Black *
    Hispanic, data = AllRaces.df)

Residuals:
       1        2        3        4        5        6        7        8        9
-0.0065279 -0.0091691 0.0476255 0.0058250 -0.0180826 -0.0249978  0.0039722  0.0010773  0.0002773

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)    3.167134   7.486517   0.423    0.713
White         -0.292915   1.206774  -0.243    0.831
Black         -0.029543   0.376870  -0.078    0.945
Hispanic       0.431258   0.485781   0.888    0.468
```

```
White:Black     0.040593   0.038443   1.056    0.402
White:Hispanic -0.004116   0.018510  -0.222    0.845
Black:Hispanic -0.011594   0.015050  -0.770    0.522

Residual standard error: 0.04122 on 2 degrees of freedom
Multiple R-Squared: 0.9998,     Adjusted R-squared: 0.999
F-statistic:  1401 on 6 and 2 DF,  p-value: 0.0007132
```

Again, all the P-values are not statistically significant for each of the explanatory variables, including the interaction terms, so it makes no sense to consider that model.

The final model is that All Races Poverty Level = .009684 + 1.005638*White + .1041087*Black + -.022477* Hispanic, in which the explanatory variables are the poverty percentages of each race per each year of peacetime between the Vietnam War and the Gulf War.

For the final poverty analysis, we now consider the Gulf War, in which the Asian poverty level is another explanatory variable to be considered and any interactions between the other races as well is to be considered. Let's first look at the no-interaction case.

```
> summary(allraces.reg)
Call:
lm(formula = AllRaces ~ White + Black + Hispanic + Asian, data = AllRaces.df)
Residuals:
       1        2        3        4        5        6        7        8
 0.010886 -0.025254  0.018386  0.015002 -0.032785  0.009576 -0.012974  0.017165
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.342712   0.532960   0.643 0.565972
White       0.894487   0.059571  15.015 0.000641 ***
Black       0.070872   0.015277   4.639 0.018874 *
Hispanic    0.038671   0.027187   1.422 0.250061
Asian       0.022263   0.007706   2.889 0.063055 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.03135 on 3 degrees of freedom
Multiple R-Squared: 0.9994,     Adjusted R-squared: 0.9987
F-statistic:  1304 on 4 and 3 DF,  p-value: 3.446e-05
```

Again, the P-value is very small; it is 3.446×10^{-5}. This P-value reveals statistical significance in showing that the level of poverty on the national level is highly dependent on the poverty levels of Caucasians, African Americans, Hispanics, and Asians. The R-squared value is 99.94%, again, nearly 100% of the variation in poverty level is explained by the regression model in which this incorporates the changes in poverty levels of each of the four races. Based on the regression analysis, it seems that the poverty level of Caucasians is going to predominantly influence the poverty level on the national level during the Gulf War. This will be followed by the poverty level of African Americans, then Asians, then Hispanics, simply in order of the P-values associated with each of these explanatory variables. The equation for the regression line as of now (without doing a stepwise regression) is All Races Poverty Level = .342712 + .894487*White + .070872*Black + .038671*Hispanic + .022263*Asian. Now, let's check stepwise regression and see if we can get rid of any of the explanatory variables. Perhaps one of them does not explain much of the variation and is insignificant to the overall representation of poverty on the national level in this country throughout the time of the Gulf War. Going through the stepwise regression, we see that All Races Poverty Level = .342712 + .894487*White + .070872*Black + .022263*Asian + .038671*Hispanic, in this order of terms. This suggests that the level of Asian poverty

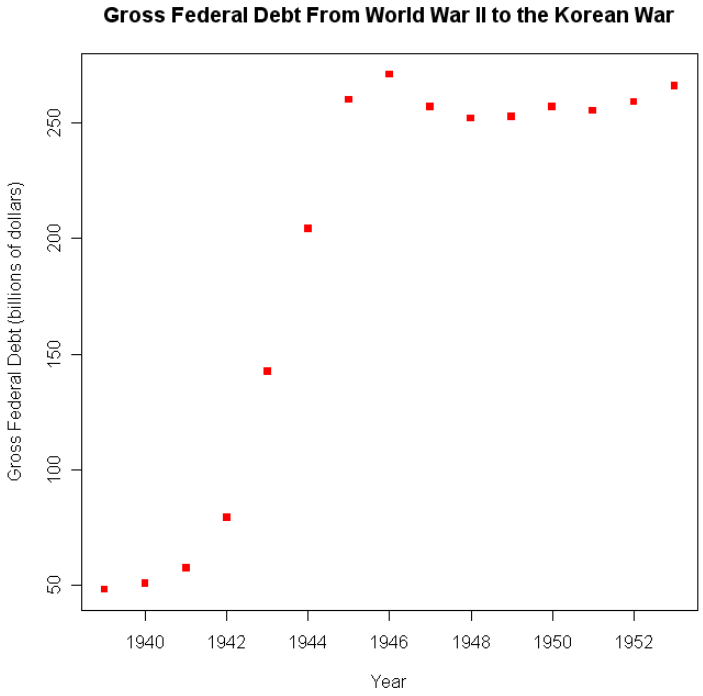
carries more weight in influencing the overall poverty level of this nation. We must recall the politics of the times; the American markets were just about to enter globalization. This explains based on the political aspects why the Asian poverty level carries more precedence than the Hispanic poverty level. With too few observations, it is insignificant to check for cross term interactions.

Question: Did the wars bring us out of deficit, give us a surplus, or plunge us into a bigger one than we already had?

One of the questions sought out to answer is about Federal Debt during the wars, what is the geometric or cyclical pattern during war and peacetime? Another question to answer is what factors or what explanatory variables attribute to the decline or increase in debt in both war and peacetime. In this statistical analysis, four wars and the peacetime in between wars shall be examined. The following explanatory variables will be considered throughout the report:

GDP
GNP
Exports
Imports
Unemployment percentage
National Defense expenditures
Surplus – Deficit
Gross Private Domestic Investment

The first period of time to be examined is the lump sum of World War II and the Korean War as they start right after each other without much time in between. Let's first inspect a graph of Federal Debut during this period of time and examine the changes.



As one can see, the Federal Debt is climbing upwards right until the end of the war, and then with the small period of peacetime in between the wars, the debt decreases, followed by another increase of debt which signifies the next war, which in this case is the Korean War. Visually, one can see that the debt in World War II followed an exponential model, while the debt in the Korean War timeframe followed a parabolic structure.

The more important thing to analyze is what were the explanatory variables that determined the shaping of the Federal Debt during this time frame. Also, how were the explanatory variables related to each other? When we do a regression analysis of the model that has every explanatory variable, we get the following results shown below:

```
summary(FedDebt.reg)
Call:
lm(formula = GrossFedDebt ~ GDP + GNP + Exports + Imports + unemployment +
   NationalDefense + SurpDef + Investment)
Residuals:
   Min     1Q  Median     3Q    Max
-29.362 -13.661  -3.853   9.071  37.750
Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)    -300.881    205.687  -1.463   0.1938
GDP             166.086    120.695   1.376   0.2180
GNP            -164.119    120.559  -1.361   0.2223
Exports          14.089      5.823   2.420   0.0519 .
Imports         -27.520     39.031  -0.705   0.5072
unemployment     12.563     11.664   1.077   0.3228
NationalDefense   1.532      2.933   0.522   0.6203
SurpDef          -1.269      2.082  -0.610   0.5645
Investment        4.842      5.642   0.858   0.4237
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 29.26 on 6 degrees of freedom
Multiple R-Squared: 0.9552,     Adjusted R-squared: 0.8954
F-statistic: 15.97 on 8 and 6 DF,  p-value: 0.001628
```

The R-squared value is very high with 95.52% of the variation in Federal Debt throughout this time period ix explained by the overall changes in all the explanatory variables put together. The P-value further supports this claim since the p-value is .001628. This P-value is statistically significant. The next thing that we have to analyze is the correlation matrix between all our explanatory variables. Below is the particular given correlation matrix:

Correlation Matrix of Variables From World War II to Korean War

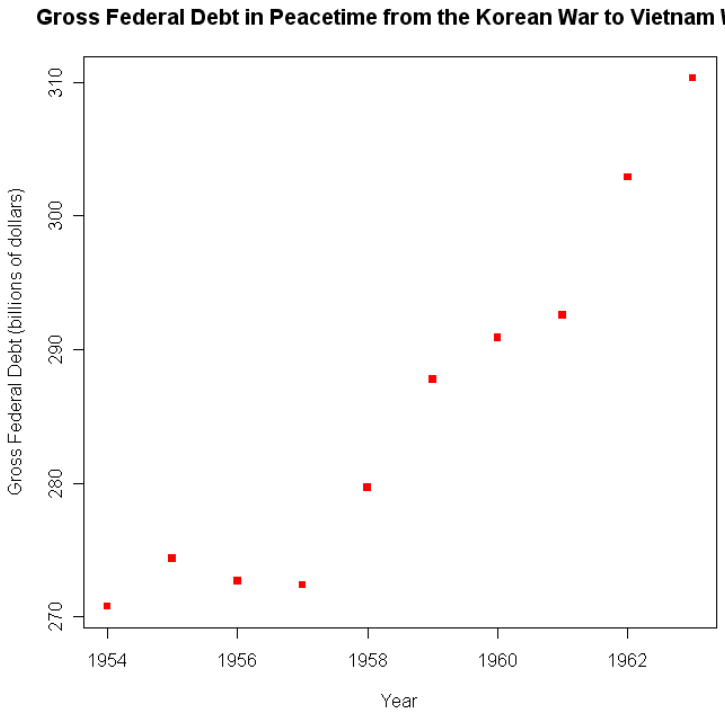
```
> cor(X)
```

	GDP	GNP	Exports	Imports	unemployment	NationalDefense	GrossFedDebt	SurpDef	Investment
GDP	1.0000000	0.9999942	0.7927909	0.9779773	-0.6896030	0.2836540	0.85651319	0.22181359	0.8517142
GNP	0.9999942	1.0000000	0.7939060	0.9784754	-0.6874658	0.2808627	0.85571926	0.22488826	0.8532680
Exports	0.7927909	0.7939060	1.0000000	0.7648778	-0.4051977	-0.2273298	0.79967248	0.61513207	0.8838685
Imports	0.9779773	0.9784754	0.7648778	1.0000000	-0.5587667	0.2003617	0.76309920	0.31191570	0.8917997
unemployment	-0.6896030	-0.6874658	-0.4051977	-0.5587667	1.0000000	-0.7233357	-0.74018467	0.38025988	-0.3090314
NationalDefense	0.2836540	0.2808627	-0.2273298	0.2003617	-0.7233357	1.0000000	0.26095179	-0.82714560	-0.2349249
GrossFedDebt	0.8565132	0.8557193	0.7996725	0.7630992	-0.7401847	0.2609518	1.00000000	0.09078923	0.6838388
SurpDef	0.2218136	0.2248883	0.6151321	0.3119157	0.3802599	-0.8271456	0.09078923	1.00000000	0.6645315
Investment	0.8517142	0.8532680	0.8838685	0.8917997	-0.3090314	-0.2349249	0.68383883	0.66453148	1.0000000

As one can see, if we focus on Gross Federal Debt, we can see that Federal Debt is strongly correlated with GDP, GNP, Export Volume, Import Volume, and Gross Private Domestic Investment. Gross Federal Debt is highly negatively correlated with unemployment. Due to the strong positive correlations, this is further evidence that as these explanatory variables increase by the billions of dollars

so does the Gross Federal Debt Despite the fact that the debt is growing, the economy is healthy because there is more money to be invested, the GDP is growing as well as the GNP. Also, due to the negative correlation between Gross Federal Debt and unemployment, this serves as another good measure to show that as the economy is getting healthy, and even though the Federal Debt looms large, the unemployment in this country during this time period was decreasing. More people are employed, thus making more money, and thus per capita income will increase as well. As one can see, the period of time between World War II and the Korean War had a sky-rocketing Federal Debt, but the economy of the United States did improve drastically based on the relationships between Federal Debt and GNP, GDP, unemployment, Gross Private Domestic Investment, Import and Export Volume.

The next thing to analyze is what happens in the years of peace between the Korean War and the Vietnam War. The years in question here will be 1954 to 1963. We will do the same modeling and same format of analysis as was done for the time period of World War II to the conclusion of the Korean War.



In peacetime, the federal debt rose tremendously and even sky-rocketed more as one can see.

Below is the regression analysis with all the explanatory variables thrown into the model.

```
summary(FedDebt.reg)
Call:
lm(formula = GrossFedDebt ~ GDP + GNP + Exports + Imports + unemployment +
   NationalDefense + SurpDef + Investment)
Residuals:
      1       2       3       4       5       6       7       8       9      10
-0.08135  0.14627 -0.04332 -0.10312  0.04166 -0.15587  0.15269 -0.01351  0.36378 -0.30722
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  183.26807   26.78352   6.843   0.0924 .
GDP            3.49697    5.06528   0.690   0.6153
GNP           -2.67183    5.06811  -0.527   0.6911
```

```
Exports         -4.69043    0.54091  -8.671   0.0731 .
Imports         -5.07580    1.72831  -2.937   0.2089
unemployment    -5.32702    1.00301  -5.311   0.1185
NationalDefense -0.06658    0.72480  -0.092   0.9417
SurpDef          0.90715    0.17048   5.321   0.1183
Investment      -0.65973    0.47216  -1.397   0.3955
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5628 on 1 degrees of freedom
Multiple R-Squared: 0.9998,     Adjusted R-squared: 0.9983
F-statistic: 675.1 on 8 and 1 DF,  p-value: 0.02976
```

As one can see, this model is perfect fit. 99.98% of the variation in Federal Debt is explained by the overall changes. The p-value is small by being ,02976, but not as significant as the p-value of the time period of war for World War II and the Korean War. This suggests that perhaps there is some lurking variable not present in the study not taken into effect.

Let's take a look at the correlation matrix in the peacetime between the Korean War to the Vietnam War, and understand what affects Federal Debt.

Correlation Matrix During Peacetime of Korean War to Vietnam War

```
> cor(X)
```

	GDP	GNP	Exports	Imports	unemployment	NationalDefense	GrossFedDebt	SurpDef	Investment
GDP	1.0000000	0.9999956	0.93323466	0.9908052	0.4087849	0.9456284	0.9666772	-0.42955241	0.9372411
GNP	0.9999956	1.0000000	0.93386640	0.9906379	0.4074190	0.9456654	0.9665776	-0.42817415	0.9375096
Exports	0.9332347	0.9338664	1.00000000	0.9265246	0.2367559	0.8622307	0.8398407	-0.09788362	0.8899036
Imports	0.9908052	0.9906379	0.92652460	1.0000000	0.3650753	0.9331350	0.9383400	-0.43365448	0.9433894
unemployment	0.4087849	0.4074190	0.23675593	0.3650753	1.0000000	0.5400396	0.4513544	-0.42842798	0.1034804
NationalDefense	0.9456284	0.9456654	0.86223069	0.9331350	0.5400396	1.0000000	0.9142038	-0.47542252	0.8030750
GrossFedDebt	0.9666772	0.9665776	0.83984071	0.9383400	0.4513544	0.9142038	1.0000000	-0.52771988	0.8998399
SurpDef	-0.4295524	-0.4281741	-0.09788362	-0.4336545	-0.4284280	-0.4754225	-0.5277199	1.00000000	-0.3909843
Investment	0.9372411	0.9375096	0.88990363	0.9433894	0.1034804	0.8030750	0.8998399	-0.39098431	1.0000000

Federal Debt is strongly correlated positively with GDP, GNP, Imports, Exports, National Defense Spending, and Gross Domestic Investment. The thing that is different between peacetime and wartime is the lack of strong correlation in either direction with unemployment. However, per capita income is increasing, so on the whole, the economy is also improving despite the debt and lack of full employment during peacetime.

Now, it is time to analyze the Vietnam War itself, one of the longest quagmires in the history of the United States other than for the Cold War itself. The years of the war are from 1964 to 1974. Let's take a look at the trend of Federal Debt during the Vietnam War by looking at the graph below.

The Vietnam War was really split into two parts, and right at the midpoint from 1968 to 1969 during a quiet time of the war within, the debt does decrease. This follows the behavior of the Korean War in the sense that debt did decrease right after a war. Just this decrease happened during a stalemate within the war.

Let's look at the summary of the full regression equation in the Vietnam War.

```
Call:
lm(formula = GrossFedDebt ~ GDP + GNP + Exports + Imports + unemployment +
    NationalDefense + SurpDef + Investment)
Residuals:
     1      2      3      4      5      6      7      8      9     10     11
 1.8512 -1.1182 -0.3356 -1.1282  0.8432 -0.5931  1.8140 -3.6120  2.2192  0.4712 -0.4118

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)    162.6242    53.8275   3.021   0.0943 .
GDP            -11.2364     6.8910  -1.631   0.2446
GNP             11.5356     6.9470   1.661   0.2387
Exports         -1.4830     1.3762  -1.078   0.3939
Imports          0.1003     0.7283   0.138   0.9031
unemployment     0.5214     6.9690   0.075   0.9472
NationalDefense -0.2093     0.9282  -0.225   0.8426
SurpDef         -0.3288     0.3064  -1.073   0.3955
Investment      -0.4851     0.6204  -0.782   0.5162
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 3.793 on 2 degrees of freedom
Multiple R-Squared: 0.9992, Adjusted R-squared: 0.9958
F-statistic: 295.1 on 8 and 2 DF, p-value: 0.003381

Again, the P-value is very small and the R-squared value is very high; so this model clearly shows that Federal Debt strongly depends on all of the explanatory variables listed in the model.

Let's look at the correlation Matrix during the time of the Vietnam War:

Correlation Matrix During Vietnam War

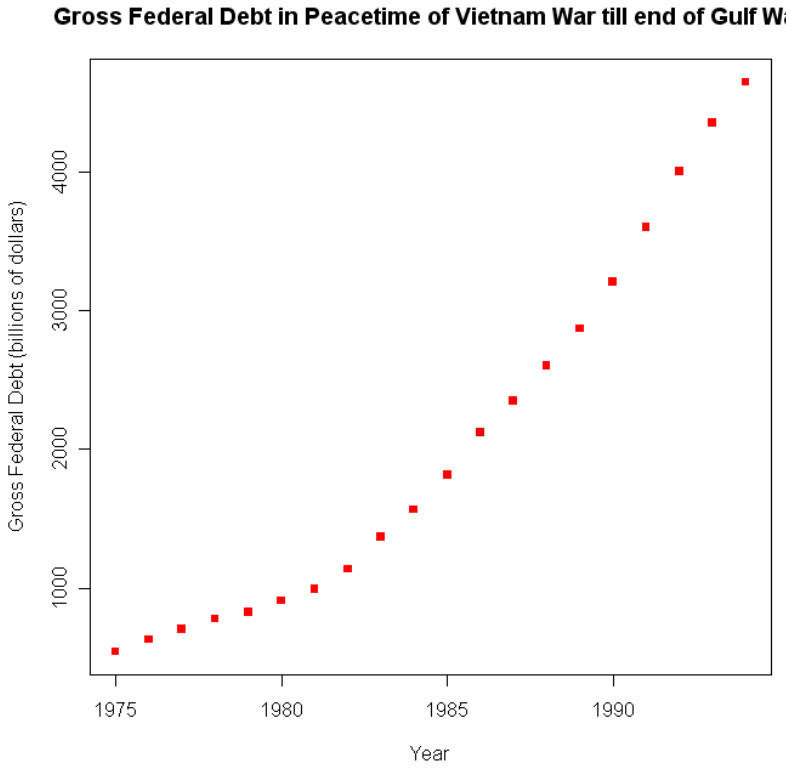
```
> cor(X)
```

	GDP	GNP	Exports	Imports	unemployment	NationalDefense	GrossFedDebt	SurpDef	Investment
GDP	1.0000000	0.9999883	0.9430317	0.9648227	0.5256689	0.8200414	0.9940257	-0.3229887	0.9868606
GNP	0.9999883	1.0000000	0.9445011	0.9659146	0.5259715	0.8183520	0.9941585	-0.3215542	0.9872050
Exports	0.9430317	0.9445011	1.0000000	0.9940727	0.5142243	0.6853918	0.9372192	-0.1469143	0.9382199
Imports	0.9648227	0.9659146	0.9940727	1.0000000	0.5230487	0.7331172	0.9592767	-0.2093760	0.9560487
unemployment	0.5256689	0.5259715	0.5142243	0.5230487	1.0000000	0.1215146	0.5624020	-0.3401518	0.5054582
NationalDefense	0.8200414	0.8183520	0.6853918	0.7331172	0.1215146	1.0000000	0.7922940	-0.3702893	0.7561369
GrossFedDebt	0.9940257	0.9941585	0.9372192	0.9592767	0.5624020	0.7922940	1.0000000	-0.3946355	0.9868230
SurpDef	-0.3229887	-0.3215542	-0.1469143	-0.2093760	-0.3401518	-0.3702893	-0.3946355	1.0000000	-0.3186692
Investment	0.9868606	0.9872050	0.9382199	0.9560487	0.5054582	0.7561369	0.9868230	-0.3186692	1.0000000

Let's look at how Federal Debt is correlated with all the explanatory variables during the Vietnam War. Federal Debt again is strongly positively correlated with GDP, GNP, Exports, Imports, Gross Private Domestic Investment. There is mild correlation with National Defense expenditures and unemployment.

We must realize that unemployment was not changed as significant and did not improve towards the end of the Vietnam War because of the fact that the Vietnam War had different characteristics compared to World War II and the Korean War. The World War II had a full scale mobilization and everybody was encouraged to work and be employed. World War II had a full scale economic mobilization that did bring us out of a Depression. The Vietnam War was a quagmire and during the second phase of the war, it was seen that unemployment did increase despite the overall health of the economy.

Now, it is time to look at the peacetime between the Vietnam War through the end of the Gulf War and up to three years past. Again, the Gulf War is only one year or length, so it is more beneficial to look at the whole chunk of time that has passed. The years of interest will be from 1975 to 1994. Please note that 1990-1991 is the year of the Gulf War.



As one can see, Federal Debt rises throughout from the end of the Vietnam War all the way through the Gulf War. The reason for this particular change in overall behavior of debt is that the Gulf War was too short of a war itself to affect the economy as a whole.

Let's look at the regression equation and see if our overall model is good in predicting the Federal Debt with all of our explanatory variables listed above. Here is the summary of the regression summary call:

```
> summary(FedDebt.reg)
Call:
lm(formula = GrossFedDebt ~ GDP + GNP + Exports + Imports + unemployment +
   NationalDefense + SurpDef + Investment)
Residuals:
   Min     1Q  Median     3Q    Max
-147.61  -20.79   13.17   37.43   86.44

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    -209.34756  236.52735  -0.885 0.395045
GDP              -0.05286    0.03174  -1.666 0.123999
GNP               1.88662    0.36799   5.127 0.000330 ***
Exports          -5.41295    1.15598  -4.683 0.000669 ***
Imports           3.95409    2.34406   1.687 0.119750
unemployment    -42.61963   24.86913  -1.714 0.114576
NationalDefense  -8.16607    1.16595  -7.004 2.26e-05 ***
SurpDef          -0.64793    0.90001  -0.720 0.486588
Investment       -3.99845    0.69170  -5.781 0.000123 ***
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 75.03 on 11 degrees of freedom
Multiple R-Squared: 0.9982, Adjusted R-squared: 0.9968
F-statistic: 750.3 on 8 and 11 DF, p-value: 6.024e-14

Again, the R-squared value is very high and the P-value is low, so this is a perfect model in explaining the change in Federal Debt through the overall changes in the explanatory variables listed in the model's description all the way above. The explanatory variables, of National Defense expenditures, Gross Private Domestic Investment, GNP, and Exports all had significant P-values. This significance shows the impact of how the U.S. entered the global economy and took center stage. These variables take much more force and have more weight in the global economy that we live in today.

Let's look at the Correlation Matrix during this time period as well.

Correlation Matrix of Peacetime after Vietnam War Through Gulf War

```
> cor(X)
```

	GDP	GNP	Exports	Imports	unemployment	NationalDefense	GrossFedDebt	SurpDef	Investment
GDP	1.0000000	0.8613355	0.7939124	0.8603099	-0.14995511	0.8798688	0.7682102	-0.74988867	0.9051915
GNP	0.8613355	1.0000000	0.9741419	0.9964356	-0.30121151	0.9459368	0.9801020	-0.85454561	0.9703093
Exports	0.7939124	0.9741419	1.0000000	0.9763287	-0.30849679	0.8615211	0.9811013	-0.78337468	0.9151282
Imports	0.8603099	0.9964356	0.9763287	1.0000000	-0.34832411	0.9298955	0.9786576	-0.81477751	0.9758232
unemployment	-0.1499551	-0.3012115	-0.3084968	-0.3483241	1.00000000	-0.2540515	-0.3220989	-0.06139389	-0.3483259
NationalDefense	0.8798688	0.9459368	0.8615211	0.9298955	-0.25405149	1.0000000	0.8807465	-0.88541107	0.9378245
GrossFedDebt	0.7682102	0.9801020	0.9811013	0.9786576	-0.32209887	0.8807465	1.0000000	-0.82388938	0.9211643
SurpDef	-0.7498887	-0.8545456	-0.7833747	-0.8147775	-0.06139389	-0.8854111	-0.8238894	1.00000000	-0.7960918
Investment	0.9051915	0.9703093	0.9151282	0.9758232	-0.34832588	0.9378245	0.9211643	-0.79609182	1.0000000

Again, Federal Debt is strongly correlated with GNP, Exports, Imports, National Defense Spending, and Gross Private Domestic Investment. GDP is of moderate positive correlation with Federal Debt.

What is the story in the Gulf War telling us through this correlation matrix. Knowing that GNP and per capita income is increasing as has been presented through this report already, plus with the combination of Gross Private Domestic Investment, GDP, GNP, and National Defense expenditures increasing along with Debt, the American economy was strong. The war did little to phase the positive growth. If the War had lasted longer, the economy would have adjusted more and fluctuated and different relationships would have happened between the explanatory variables in the model.

With this said, the overall conclusion from this whole project is that war is profitable despite a huge debt that has incurred through war and even during peacetime. War is thus healthy for the economy.

What about the overall model, what is the general trend of Debt, and how is it most strongly influenced? To answer this question, let's look at the summary regression call for the whole time frame from 1939 to 1994, from World War II to the end of the Gulf War.

Call:
lm(formula = GrossFedDebt ~ GDP + GNP + Exports + Imports + unemployment +
NationalDefense + SurpDef + Investment)
Residuals:
Min 1Q Median 3Q Max
-211.36 -45.52 15.41 64.03 138.38
Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 509.71477 58.05391 8.780 1.79e-11 ***

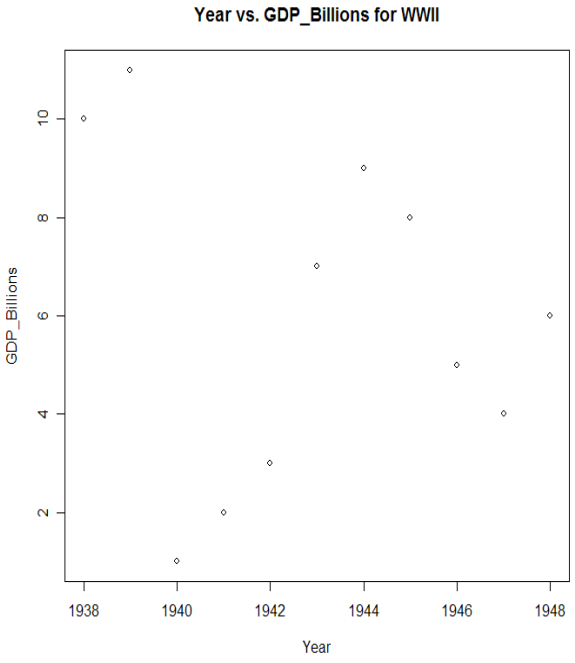
```
GDP            -0.08782    0.03446  -2.548   0.0142 *
GNP             1.34561    0.18266   7.367 2.28e-09 ***
Exports        -6.51343    1.15205  -5.654 8.96e-07 ***
Imports        10.78121    0.97378  11.072 1.08e-14 ***
unemployment   -35.31119    6.04770  -5.839 4.72e-07 ***
NationalDefense  -6.40929    1.00878  -6.353 7.84e-08 ***
SurpDef         -0.93084    0.57896  -1.608   0.1146
Investment      -6.01500    0.71503  -8.412 6.22e-11 ***
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 93.95 on 47 degrees of freedom
Multiple R-Squared: 0.9944,     Adjusted R-squared: 0.9935
F-statistic:  1051 on 8 and 47 DF,  p-value: < 2.2e-16
```

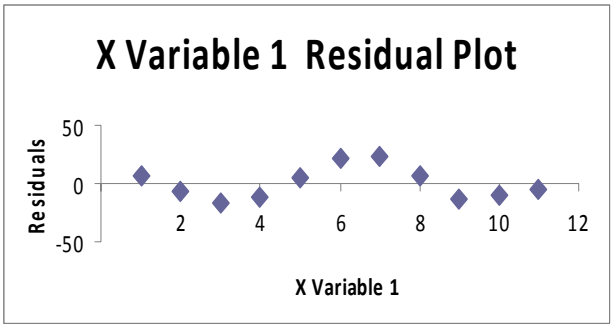
As one can see, the R-squared value again is very high and the p-value is so small that it is not even measurable. This clearly shows that all of the explanatory variables put together explain Federal Debt extremely well. What are the variables that influence Federal Debt the most. In order from highest to lowest of statistical strength and significance, Gross Private Domestic Investment, Imports, GNP, National Defense expenditures, Exports, and Unemployment level of the nation are the strongest in influencing Federal Debt.

Question: Is war profitable and beneficial to the American Economy?

We will try to analyze if war does affect the GDP and thus the overall economic situation of the US. It will show if the effect is positive or negative which is really GDP growth or decline. Please remember, as stated above that the GDP is fixed for inflation using the CPI as suggest from the presentation given. This data is also taken in quarters which mean four quarters per year so 3 years prior and after the war give us at least 24 quarters, plus the war, so we have a pretty good set of data points, except for World War II, the government did not start taking quarters until 1947.

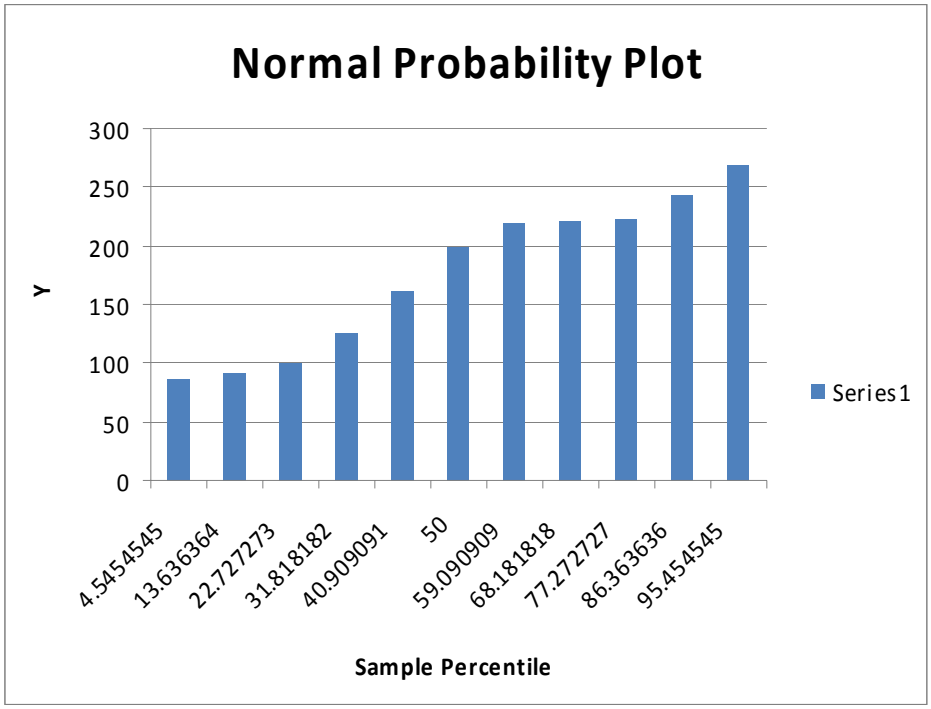
The first war that will be done is World War II and we will take the points before, during, and after, do an analysis and then take out the points of the war and decide whether this is statistically feasible to conclude that it did affect the GDP level. This will be done a little different then the other set of wars because of the limited data.





RESIDUAL OUTPUT

Observation	*Predicted Y*	*Residuals*	*Standard Residuals*
1	79.73182	6.368182	0.461113
2	99.15818	-6.95818	-0.50383
3	118.5845	-17.1845	-1.24431
4	138.0109	-11.3109	-0.81901
5	157.4373	4.462727	0.323141
6	176.8636	21.73636	1.573905
7	196.29	23.51	1.702332
8	215.7164	7.383636	0.534641
9	235.1427	-12.8427	-0.92993
10	254.5691	-10.3691	-0.75081
11	273.9955	-4.79545	-0.34723



PROBABILITY OUTPUT

Percentile	*Y*
4.545454545	86.1
13.63636364	92.2
22.72727273	101.4
31.81818182	126.7
40.90909091	161.9
50	198.6
59.09090909	219.8
68.18181818	222.3

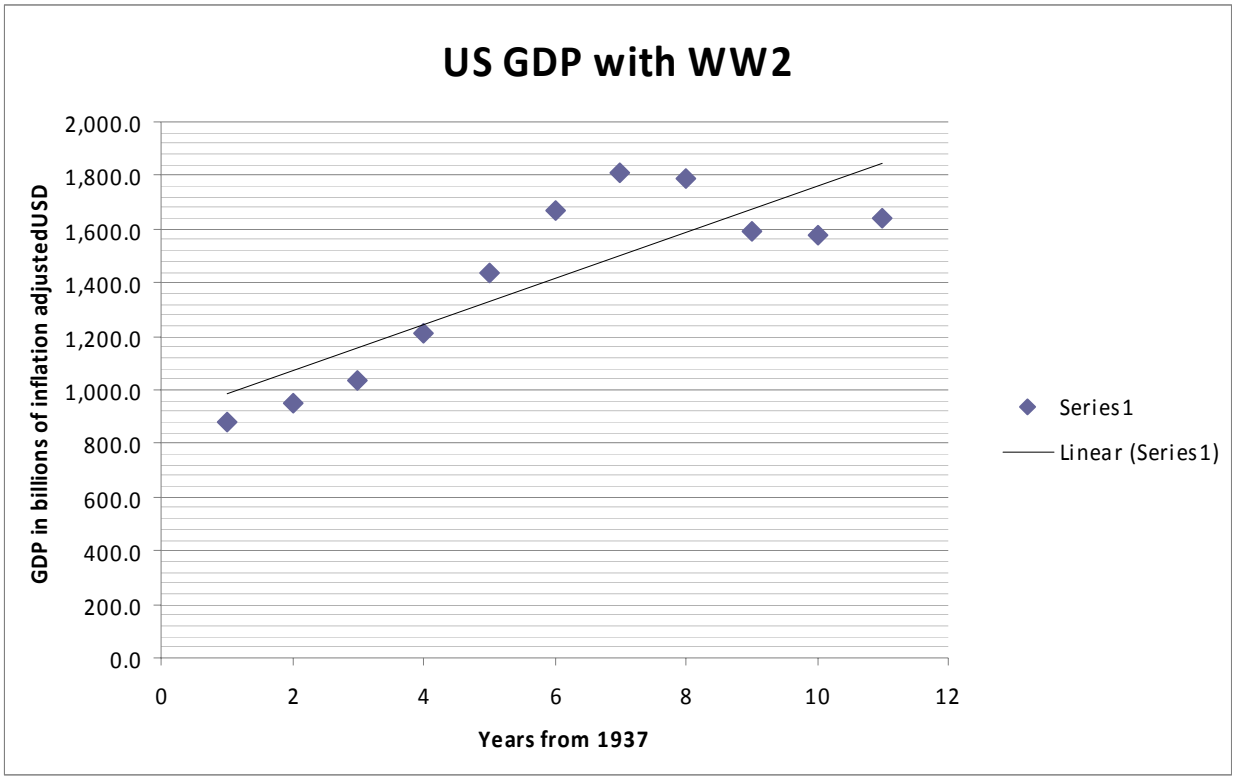
77.27272727	223.1
86.36363636	244.2
95.45454545	269.2

Regression Statistics	
Multiple R	0.97779
R Square	0.956073
Adjusted R Square	0.951192
Standard Error	14.55751
Observations	11

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	41512.2	41512.2	195.8852	2.05603E-07
Residual	9	1907.289	211.921		
Total	10	43419.49			

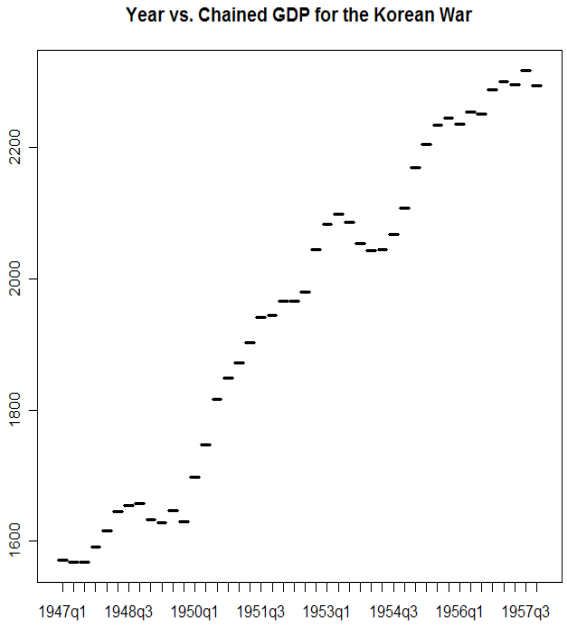
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-37568.6	2696.895	-13.9303	2.14E-07	-43669.3611	-31467.8	-43669.4	-31467.8
X Variable 1	19.42636	1.388004	13.9959	2.06E-07	16.2864809	22.56625	16.28648	22.56625

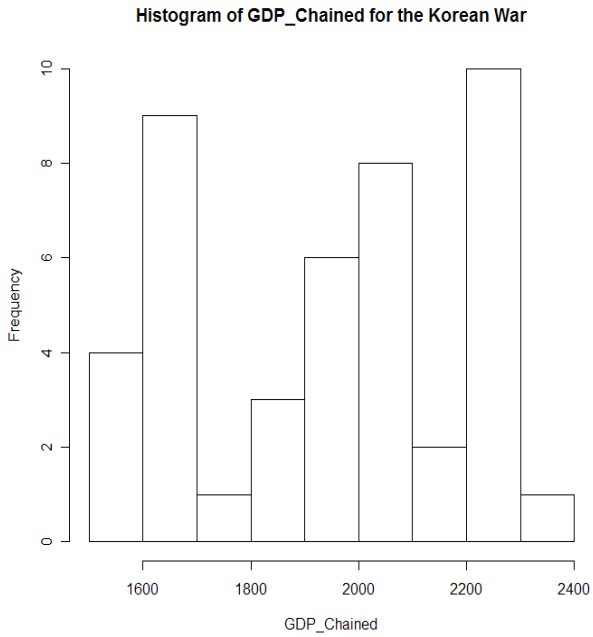


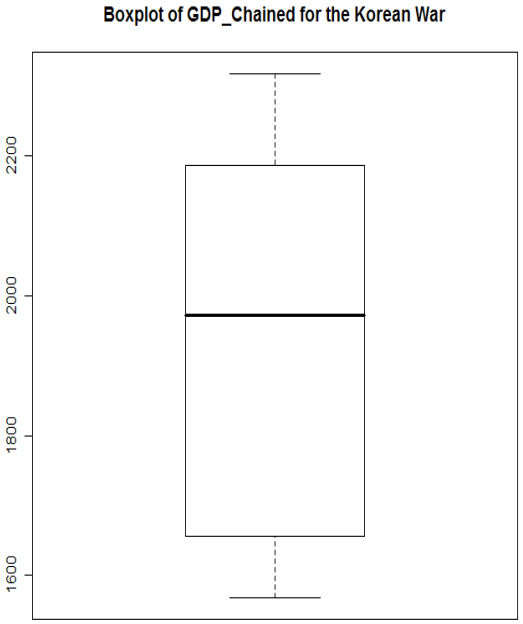
Through these calculations along with the plot of the years we see that the R^2 value is very good and the P value is greatly significant. We can conclude even though when we first went into World War

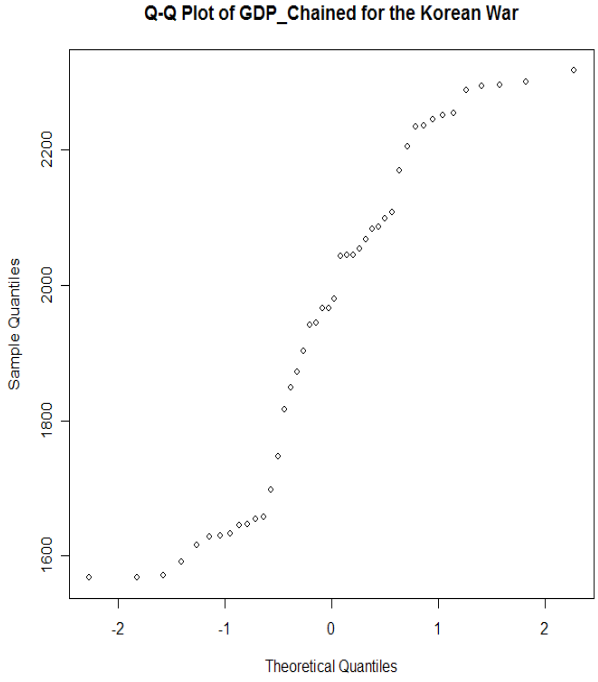
II, our GDP decreased greatly, we still came out of it because of the war and ended up with a greater GDP, thus confirming that this war did help our economy and was very profitable.

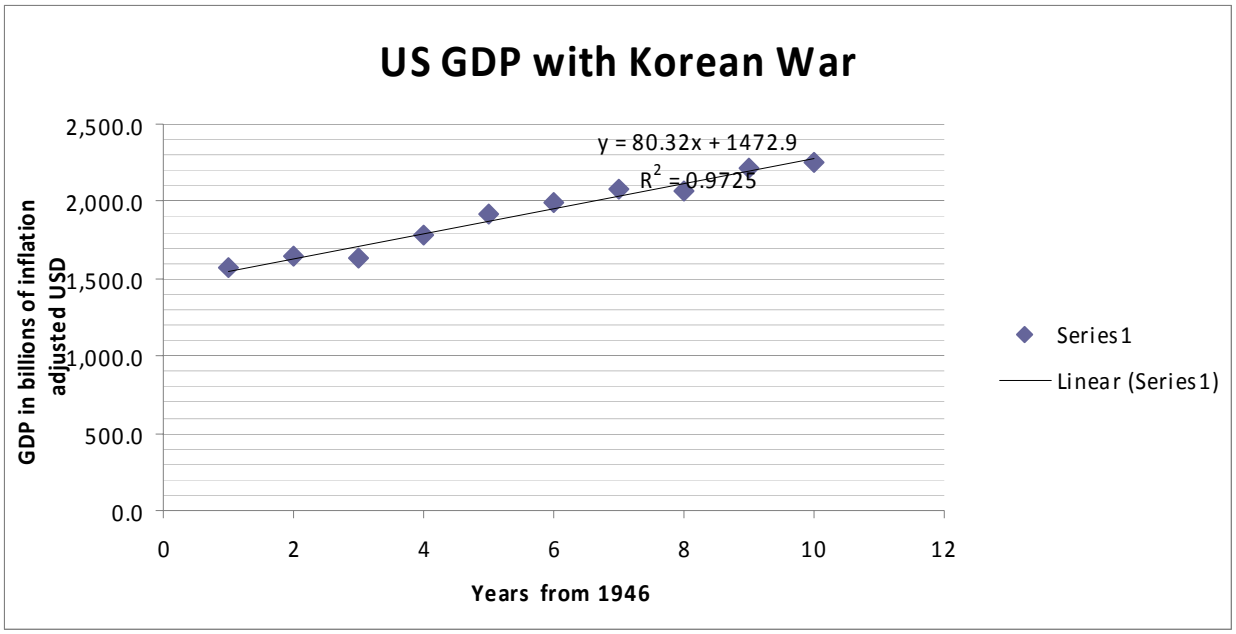
The next war that will be done is Korean War which will be done a little differently because we have all the quarters of data.





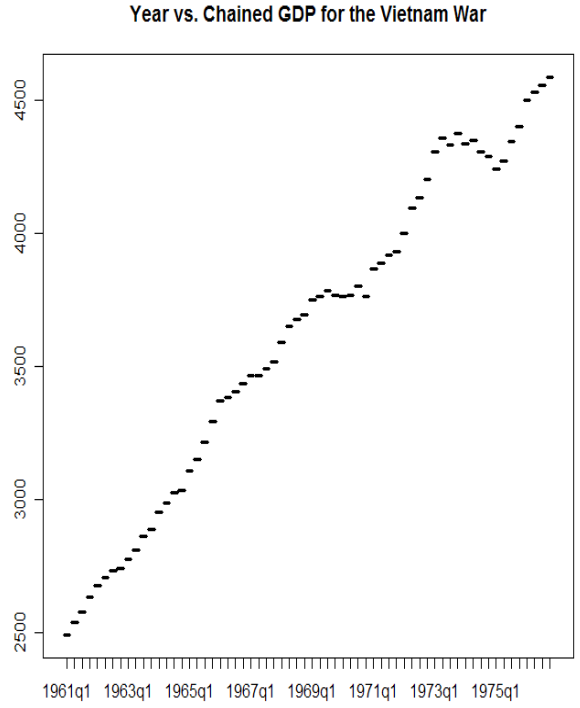


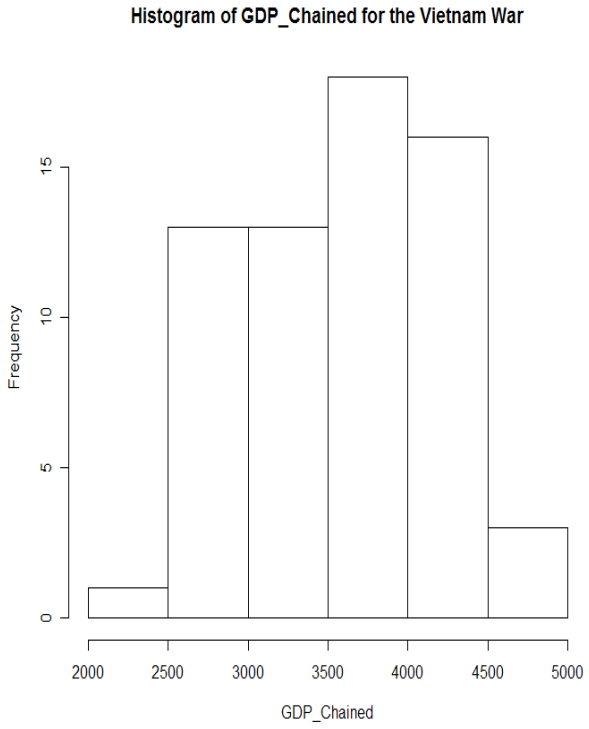


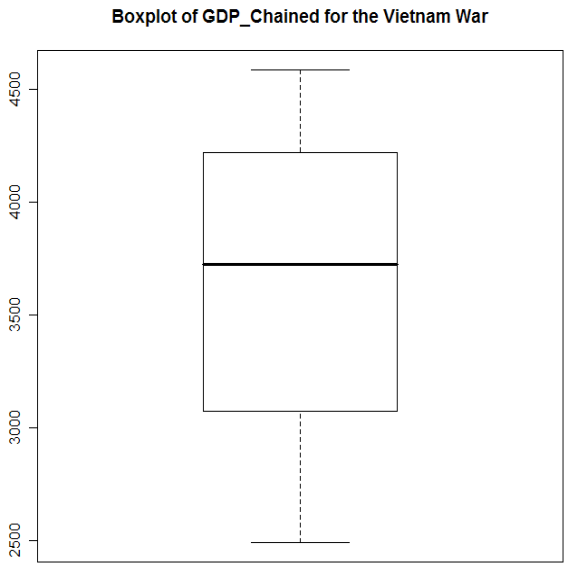


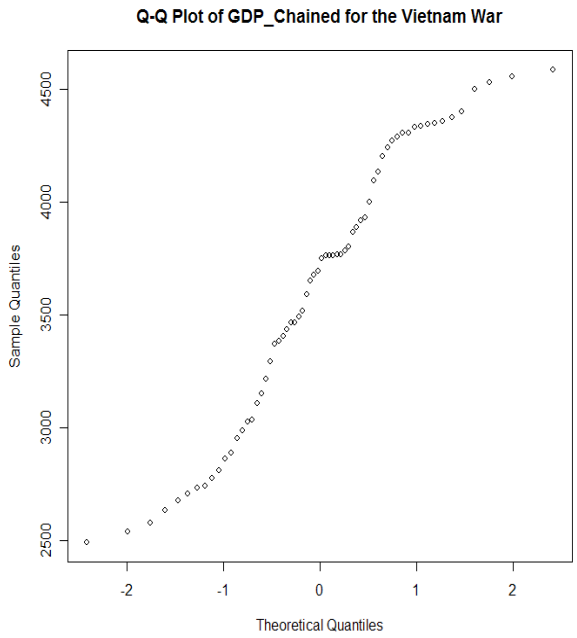
As you can see by these four plots, the Korean War helped the US out quite a bit. With the war the outcome is that the GDP rose and the US became profitable. The histogram does not seem to have a defined skew-ness but overall deciphering from the QQ plot and the main plot we see that the GDP did raise and cap, which is exactly what we saw in income for the same war. Some statistical analysis is that the R^2 value is 0.9725 which is very good and the equation for the best fit regression line is 80.32x+1472.9 which means that this war was very profitable.

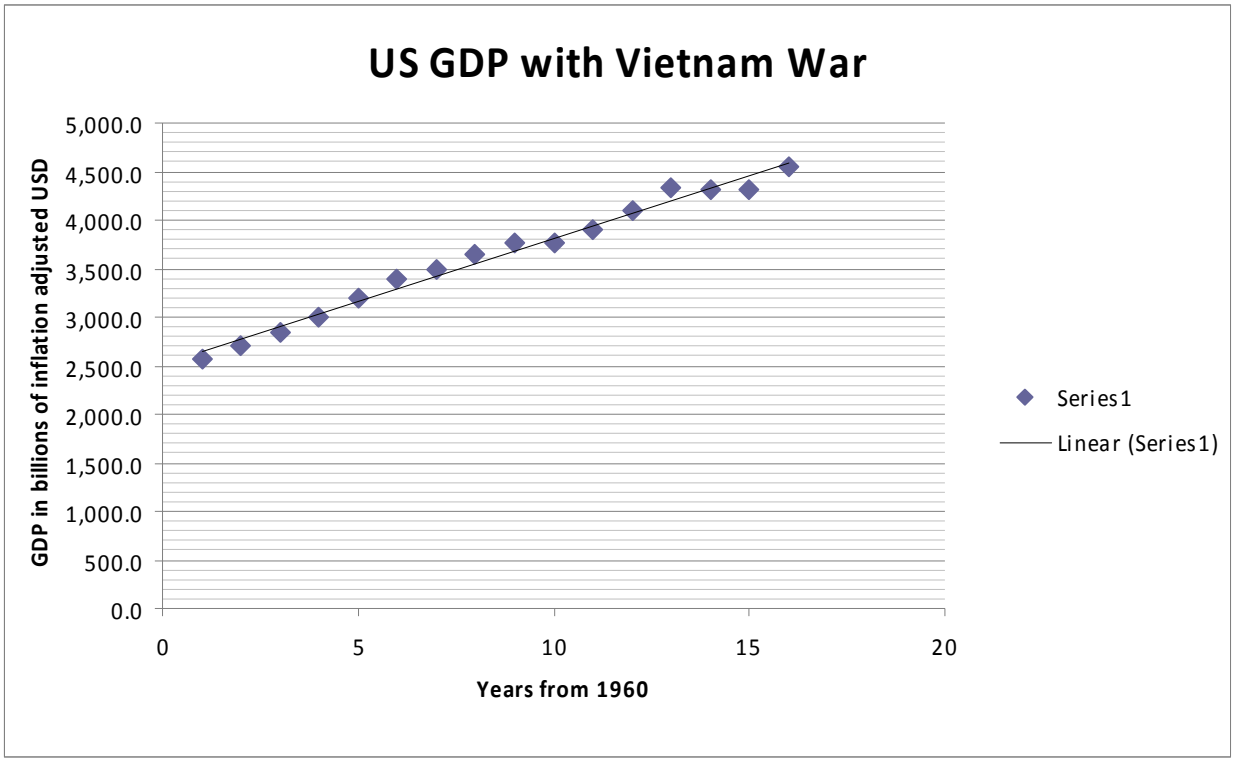
Next, we will venture into the Vietnam War and will do the same data and graphical analysis to show that this war is profitable.





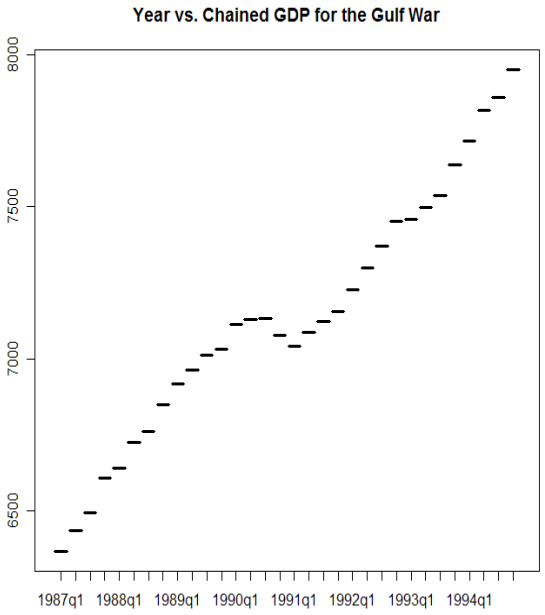


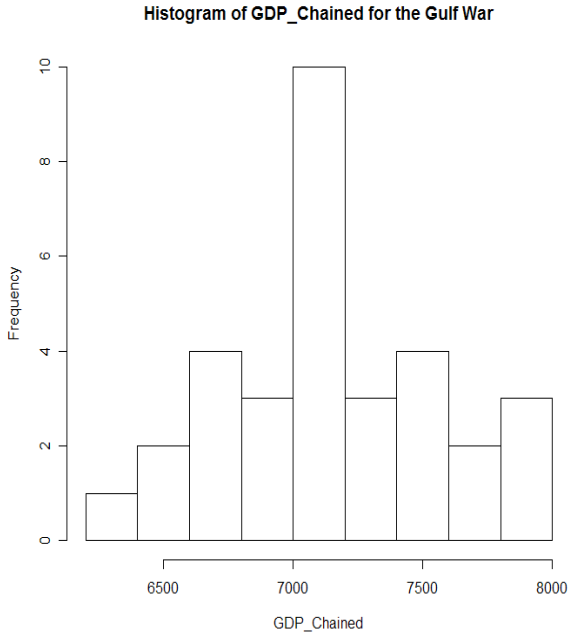


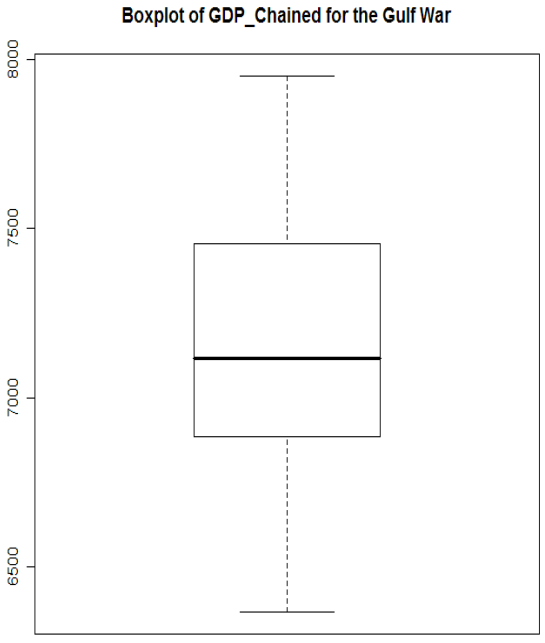


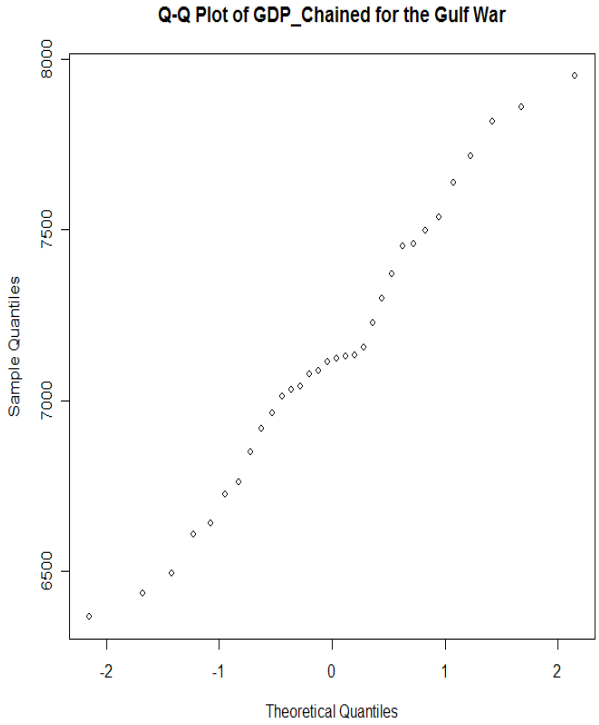
This was has a very nice linear regression with GDP almost always increasing even with inflation. The R^2 value is 0.9833 which is very good and the equation for this model is y=130.53x+2508.6. The histogram looks to be quite linear with the hill curve exactly what it should be and based on the regression plot and the regular R plot we see that GDP is always increasing and eventually caps, just like it did with income. This war though is kind of special because it is the first war that the US had actually lost and even with this fact at hand, the US still rose above and with inflation, had the GDP rise to prove that this war was profitable.

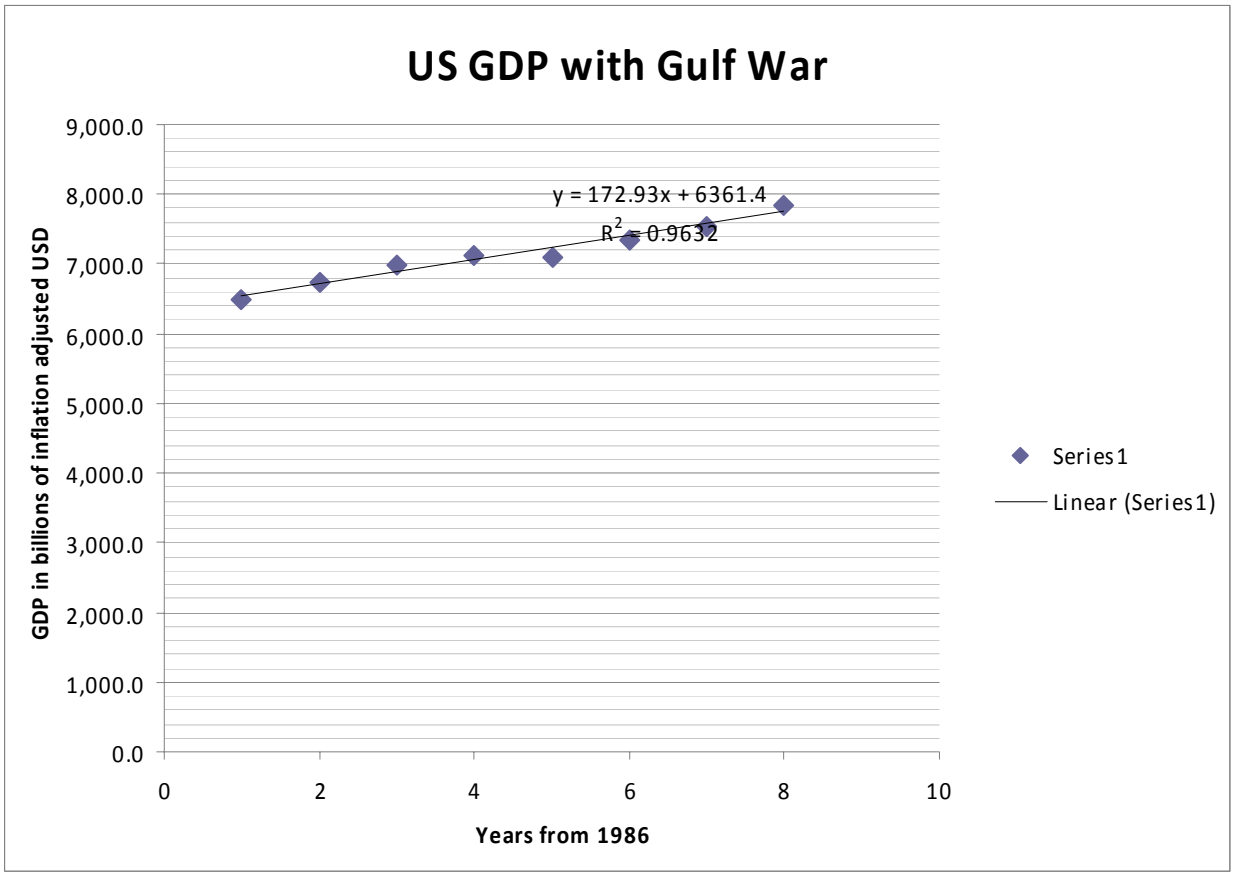
Finally, we will analyze the Gulf War with the graphical and mathematical summaries. This was a very short war to say the least. I do not think it even last more than one year, so the quarters of data will show a better detailed description of what kind of statistics is going on at this point in time.











This war based on the regression analysis with R^2 being 0.9632 which is very good and having a linear equation of y=172.93x+6361.4 was a very good war despite the fact that it was very short. From the presentation to the report after doing the CPI we realized that it became stronger correlation. Also know that GDP shoots up after we pull and during the cold war. We see through the histogram that it is quite linear with the hill going and no sort of defined skew-ness. The regular plot, Q-Q plot and regression line show us that even though it was short it was worth it because of how it profited our economy.

Conclusion

Through the various data collected we see that the answer to the ultimate question is that war is profitable through a positive correlation, and yet puts us into deficit and shows that the white race has the most poverty. This means that we reject our null hypothesis and keep our alternative hypothesis in the most positive of light. The various R commands included in this, but are not limited to the standard plot, box plot, histogram, mean, standard deviation, summary, anova, and regression analysis. We also saw that as mentioned before, the data is very hard to obtain and the internet was not a very source tool. We had to rely on paperback rather than the world wide web. All in all though, the project was a success and we used various topics from class that gave us a better understanding of the statistical subject as a whole.