**LPI CASE STUDY**

Rob O’Dell, Andrea Perez, Will Kim, Richida Gyimah

Department of Information Systems, Arizona State University

SCM 516: Descrptv&Predctv Sup Chn Analyst

Dr. Surya Ayyalasomayajula

October 30, 2022

**Introduction**

We incorporated two outside data sets to assemble our explanatory variables in our regression model while using LPI score as our dependent variable. The main purpose of our regression model was to explore what outside factors would affect our LPI score for African countries. Our first data set came from lpi.worldbank.org where we were able to find yearly data on factors involving GDP and population. The explanatory variables extracted from this data set are as follows: GDP, population, GDP per capita growth, GDP annual growth, spending on education as a percent of GDP, and exports of goods and services as a percent of GDP. For our second data set, we chose to focus on the explanatory variable that we were most interested in; CPI, or the corruption perceptions index, published annually from transparency.org. The organization’s CPI scoreis calculated on a score of 0 through 100 with 0 being the most corrupt. This corruption score was calculated in three steps. The first step was to collect a minimum of three data sources per country that captured corruption perception from the last two years. The second step was to standardize the data by converting all observations from the data source to scores from 0 to 100. The third step was to average all available standardized scores for each country, and the result would be the CPI data we merged into our regression model. Our group theorized that public sector corruption should play a part in determining the quality of trade logistics for a given country. Therefore, we came up with the null hypothesis that the beta coefficient for our CPI variable is equal to zero. We hoped our regression would show enough evidence of correlation between LPI and CPI to reject this hypothesis. The regressions shown in later pages will prove that our results are indeed statistically significant enough to reject this null hypothesis at 1% significance level.. Finally, our group chose to focus on the following four African countries: Ghana, Nigeria, Cameroon, and Angola. This decision was made from the fact that one of our group members, Richida Gyimah, **lives in** Ghana.

**Analysis**

We chose Ghana to examine and wanted to examine three similar countries in size, location, and population for the analysis. So we chose to examine Ghana, Nigeria, Cameroon, and Angola. All four have coastlines on the Atlantic Ocean, and Cameroon and Angola are close to Ghana in population. Nigeria is located close to Ghana. Three of the four countries are in West Africa, while Angola is in Southern Africa.

We looked at the LPI scores for the countries and compared them to the average for the African countries in the LPI data in each of the last five periods the LPI was compiled (2010, 2012, 2014, 2016, and 2018). The average LPI score for all African countries sampled in the five years was about 2.45 each year. In each period, 35 and 41 African countries were included in the sample. Nigeria, Cameroon, and Ghana consistently scored just about that average (in the 2.44 to 2.8 range) in each of the five years the scores were assembled, except for Cameroon which scored below average in 2016. Angola's LPI scores were below the average of other African countries each year. A time series chart of the four countries scores can be found in figure 1.1

We initially attempted to do a regression on the four countries but did not get the results we wanted because the four countries' data was so similar and not normally distributed. So we turned to examine what outside factors affected the scores by incorporating other data.

First, we turned to lpi.worldbank.org to find some of the data sets that we needed. We chose to include: GDP, GDP annual growth, GDP per capita growth, population, spending on education as a percent of GDP, and exports of goods and services as a percent of GDP. For corruption, we downloaded datasets from transparency.org, which produces an annual "Corruption Perceptions Index" of CPI that ranks 180 countries worldwide on a zero to 100 scale on their perceived levels of public sector corruption. The ranks are from highly corrupt with a score of zero to very clean with a score of 100.

We took the datasets from lpi.worldbank.com and transparency.org and pulled them into SQL to combine the relevant years together and export them to Excel to use StatTools to run regressions on the data.

We ran a series of six regressions, two for data in 2018, two in 2016, and two in 2014. Each year we examined the effect of the seven variables on the entire LPI data set and contrasted it with regressions from African countries included in the LPI data that year. We ran the regressions backward, forward, and stepwise until we found the regression with the highest R squared, with an overall p-value of <.05, and having each variable in the regression with a p-value of less than <.05. In 2018, for the overall dataset, we found that perceived public sector corruption had an enormous effect on the overall scores, with an R squared of .6594 in regression with LPI as the dependent variable and the public corruption score as the independent variable. With a p-value of <.0001 in Fig 1.1, the regression was statistically significant. Thus, we can say that the data provide enough evidence to reject our null hypothesis that corruption had no effect on the LPI data. Overall, our regression equation that explained the most with the least, raising the R squared value the most while also increasing the adjusted R squared and having p values below .05 in Fig 1.2 included both the CPI corruption index and a country's GDP.

For African countries in 2018, we found a similar effect, although the R squared between LPI and the corruption score was not as pronounced. It had a .2709 R square with LPI as the dependent variable and the CPI corruption score as the independent variable. We believe the lower score is potentially because there are virtually no examples of high “very clean” scores in the data from African countries. The CPI scores for African countries averaged about 30 in 2018, 2016 and 2014, with a maximum score under 56 in each period. The CPI for our four countries averaged 19 for Angola, 26 for Cameroon, 27 for Nigeria and 44 for Ghana.

With a p value of .0006 in Fig 1.3, the result is statistically significant and again helps us disprove our null hypothesis that corruption has no effect on the LPI logistics scores. The best fitting regression for African countries in 2018 we used was GDP along with the CPI corruption score to generate an r squared of .4268 and an adjusted r squared of .3950 as shown in Fig 1.4. Both of the variables had a p-value of <.01, giving them statistical validity. In 2016 and 2014, we found similar R squared and adjusted R squared both for the entire dataset and for the African countries in the data set that year. Our best fitting regression for all countries in 2016 had the corruption index, GDP, exports of goods and services and population as variables. The R squared was .7026 and the adjusted r squared was .6940 in Fig 1.5

Our results for African countries in 2016 were similar. Our regression that told the most with the least variables included the CPI corruption score, GDP and also a country’s exports of goods and services. The p values of all the variables were below .03, meaning they are statistically significant in Fig 1.6. We found similar results in 2014 for both African countries Fig 1.7 and the entire dataset in and Figure 1.8. Therefore, in six separate instances we were able to disprove our null hypothesis and conclude that corruption does influence LPI scores and can help explain the scores.

**Results**

In conclusion, we used outside data to get closer to the p-Value of 0.05. We took data from 2018 and assigned the dependent variable as the LPI Score and the independent variable as CPI corruption data and GDP Score. Figure 1.1 shows a correlation between the dependent and independent variables. In figure 1.1, the R-square is 0.69, and the p-Value is 0.0001. The regression model proved our hypothesis that if an R-Square is higher than 0.5 and a p-Value is lower than 0.05 would tell us that there are correlations between the independent and dependent variables. In figure 1.3, we measured the CPI corruption scores alone, and the p-Values came out low. Although the R-Square is different from the hypothesis of 0.5 however is more than halfway with fewer data. So, the data still shows a correlation because the p-Values are low. In figure 1.4, which shows data from 2018, we reached closer to the hypothesis that there is a correlation between the dependent variable LPI Score and the CPI corruption and GDP Score, but less near to analyzing data from figures 1.5 and 1.6 in the year 2016 that has other independent variables. Figure 1.5 and 1.6 shows the best correlation between independent and dependent variables. We used the dependent variable LPI Score and the independent variable CPI Corruption, GDP, exports, and population for 2016. The R-squares met our hypothesis with this data set, and the p-Value stayed at 0.0001. Lastly, this explains that the CPI Corruption, GDP, Exports, and Population are majorly affected by LPI Score. Overall, when combining all the independent variables, it shows a correlation between multiple independent and dependent variables.

**Recommendations/Outcomes**

Our regression models turned out to be statistically significant with the low given p values for our beta coefficients. Regressing LPI as a function of CPI and GDP for 2018 gave us some of the lowest p values we’ve seen throughout all our regressions, as well as displaying an acceptable R square value of 0.69 which describes goodness of fit. However when we took this regression and applied it to only African countries instead of the entire data set, our R square was reduced drastically. Only around 42 percent of the variation in LPI scores of African countries were explained by variations in CPI and GDP for 2018. We surmise this could be because CPI scores for African countries only are not normally distributed; more African scores were closer to 0 than 100. Regressing LPI as a function of CPI, GDP, exports and population for the year 2016 showed us that 70 percent of the variation in LPI was explained by these variables. Additionally, the p value for the beta coefficient for CPI is the lowest in the regression; displaying statistical significance. We were able to find statistically significant p values and high R squared values through each of these regressions involving CPI as an explanatory variable. This confirms that we have enough evidence to reject the null hypothesis that CPI doesn’t affect LPI score. From this result, we can recommend that countries looking to improve their trade logistics should focus on improving their CPI score since their LPI score will improve consequently as well. Additionally, our group can recommend that CPI be included in the calculations for LPI score in subsequent years due to its statistical significance.

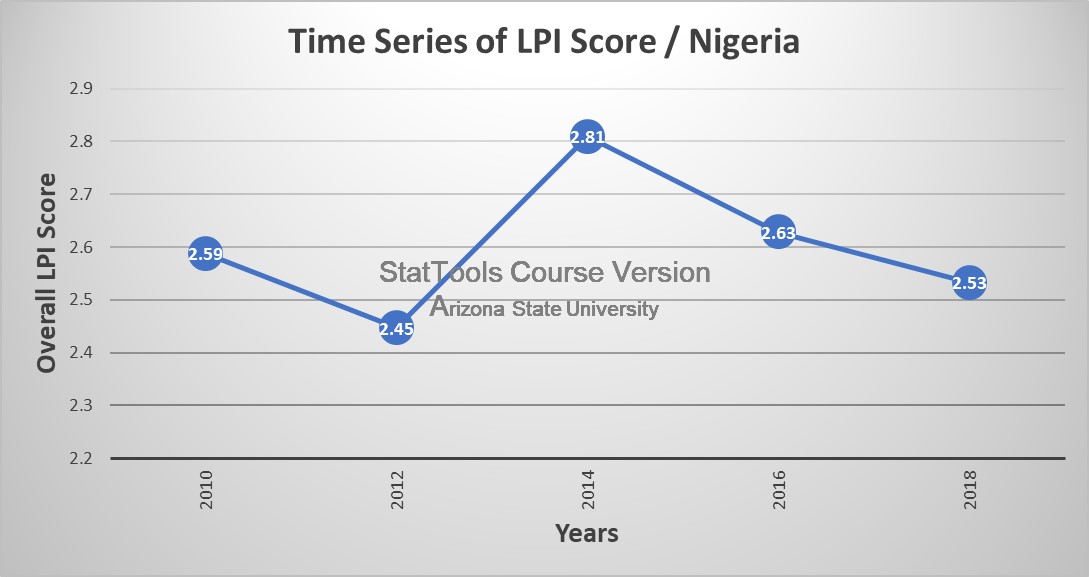
References

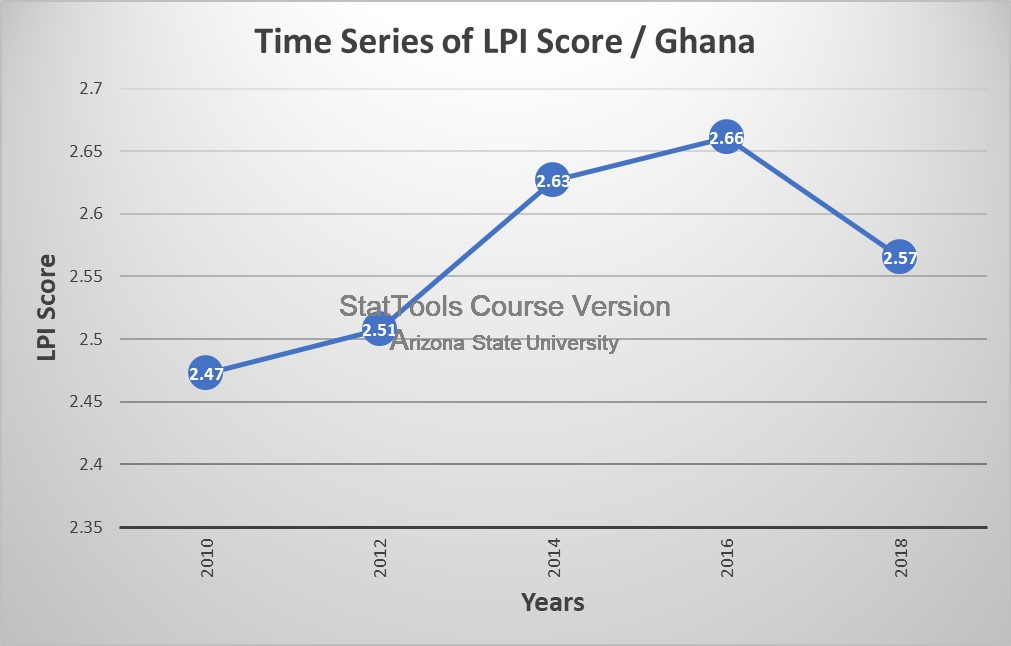
The World Bank. (2010). *Home | Logistics Performance Index*. Worldbank.org. <https://lpi.worldbank.org/>

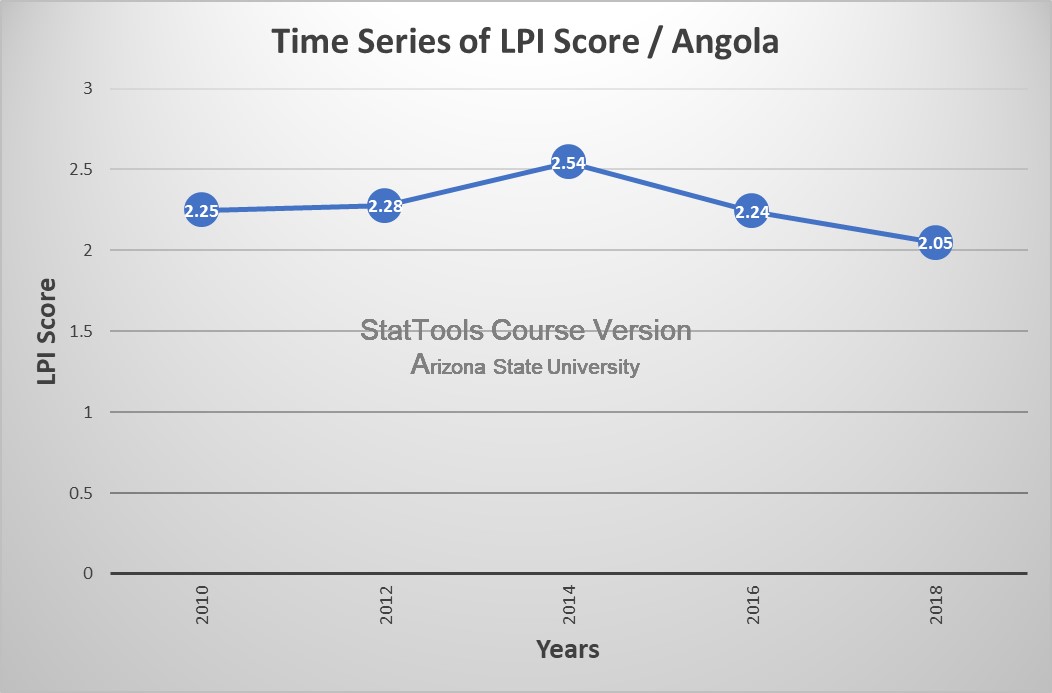
Transparency International e.V. (2000). *Transparency International - The Global*

*Anti-Corruption Coalition*. Transparency.org. <https://www.transparency.org/>

Appendices Fig 1.1







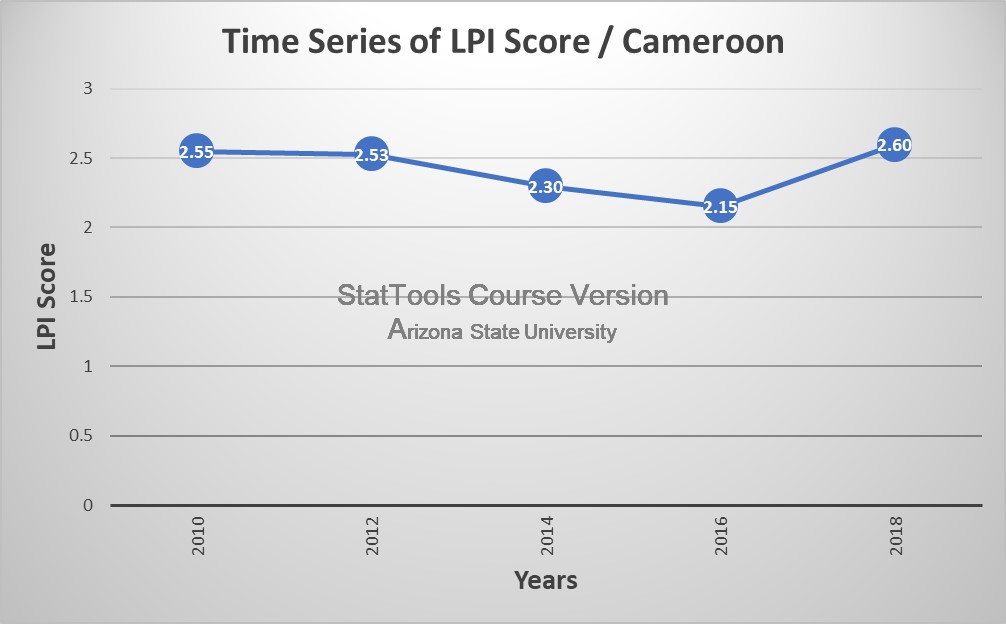


Fig 1.2

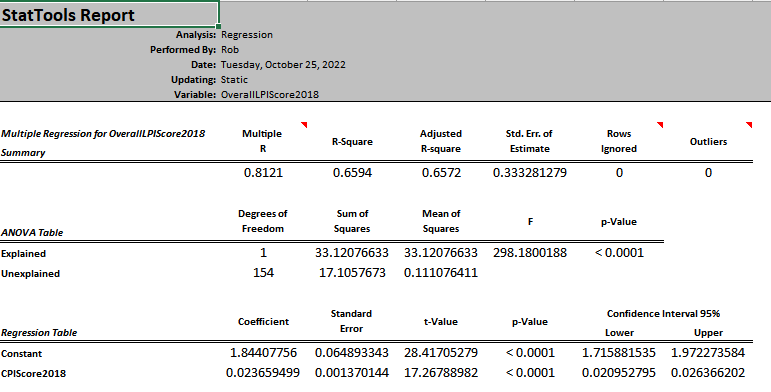


Fig 1.3

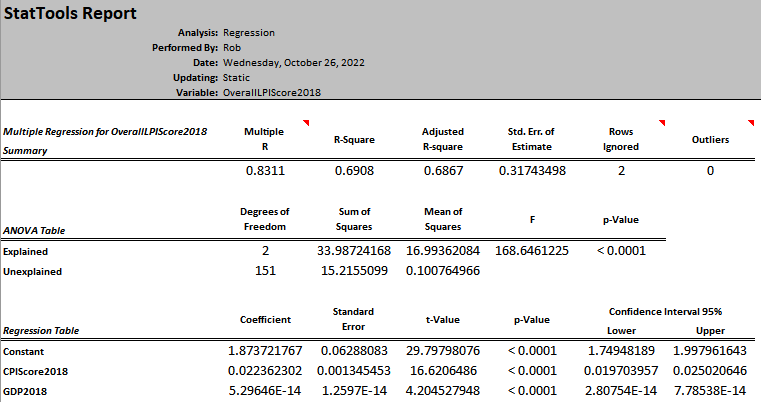


Fig 1.4

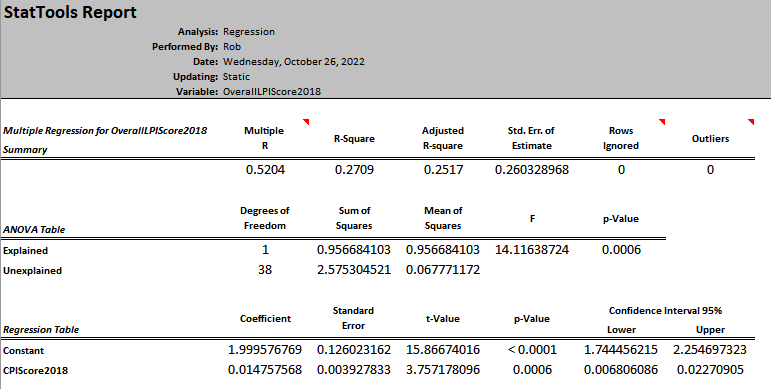


Fig 1.3

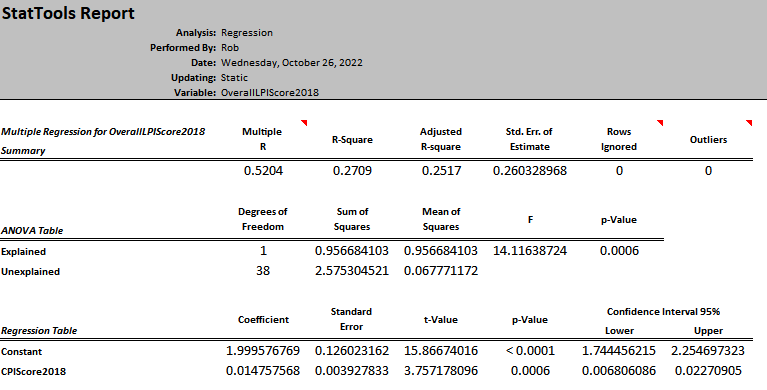


Fig 1.4

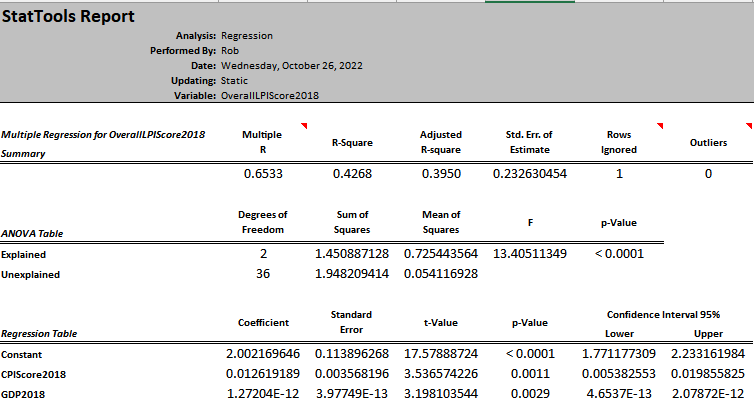


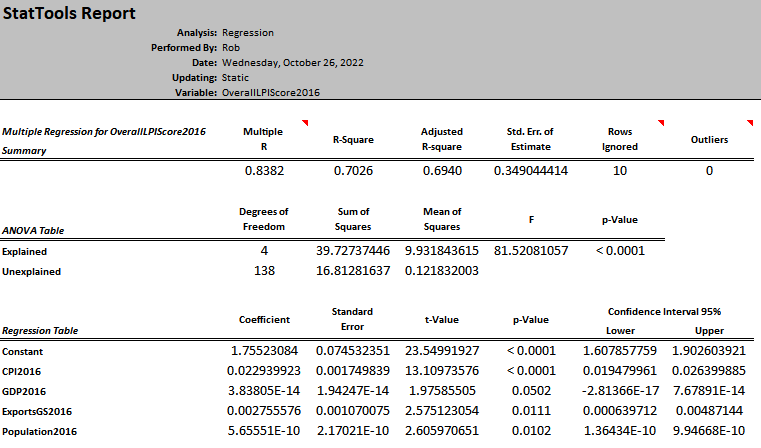
Fig 1.5

Fig 1.6

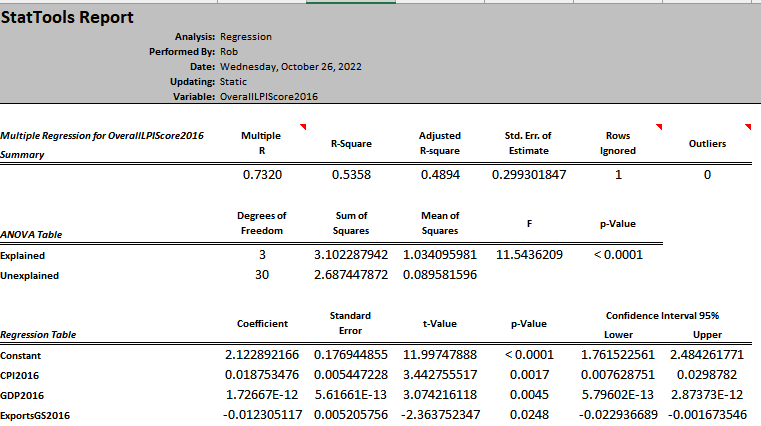


Fig 1.7

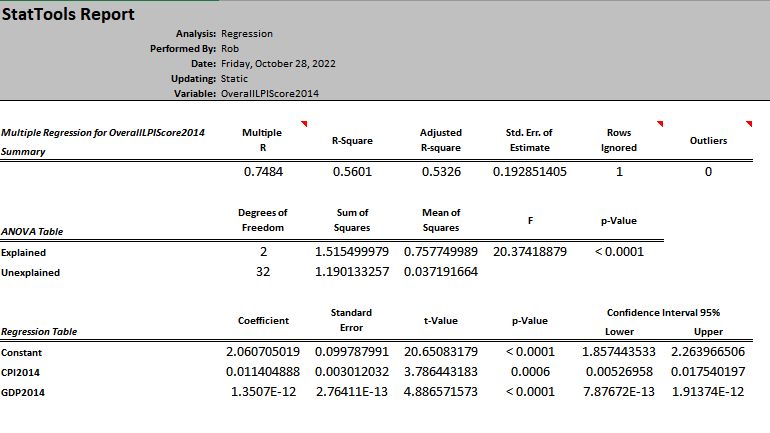


Fig 1.8

